THE LIBRARY

UNIVERSITY OF MAINE

GIFT OF

Dean L.S. Corbett

500

LIBRARY
UNIV OF MAINE

LIBRARY
UNIV OF MAINE

STATISTICS AS APPLIED IN BUSINESS

By

BENJAMIN FRANKLIN YOUNG, M.A.

Assistant Chief Accountant, New York Telephone Company; formerly Instructor in Statistics, New York University School of Commerce, Accounts and Finance

THE RONALD PRESS COMPANY
NEW YORK

Copyright, 1925, by
THE RONALD PRESS COMPANY

All Rights Reserved

PREFACE

The use of statistics has been developed into a definite business tool for controlling processes and detecting inefficiencies. In businesses where it is impossible for the executives to maintain direct contact with the various phases of the organization, some sort of statistical service has come to be essential. This volume describes the procedure to be followed in building up and directing such work.

The special difficulty in connection with a statistical service in a business is that it requires of those operating and controlling it a knowledge both of statistical technique and of the problems, practices, and routines of business management. Many of the men who have had to undertake such work have lacked definite statistical training. On the other hand, those trained as statisticians have often known too little of the concrete conditions of management. Treatises on the subject, while devoting considerable attention of late to the business cycle and other matters of broad economic nature, have dealt little with the more elementary and pressing problems of the individual business organization. This volume seeks to present the essentials of elementary statistical methods in such a way that their application to the individual concern may be readily grasped by practical business men, as well as by students of commerce preparing for a business career.

As illustrating the adjustment of the discussion to the practical requirements of the business man and the student of business, attention may be called to Chapters II to VIII, which treat in detail the question of determining the scope of the facts to be developed. From the concrete outline here given of facts regarding personnel, finances, and operations, the list of basic data for any type or size of concern may be readily determined.

In any presentation of the subject of statistics, either in the class-room or to the individual reader, one serious difficulty is that of bridging the gap between the theoretical and the practical.

206677

Frequently the student, though clearly understanding a statistical principle during its presentation, will utterly fail to recognize its application in the solution of a practical problem. The practical problems given in Appendix F are designed to remedy this defect. Some of these are simple, directly illustrating specific statistical principles. Others, more elaborate, partake of the nature of research inquiries and take the form of original statistical studies. Frequently, the practical aspect of the problem in the list is emphasized by associating with it pertinent questions on which the statistical study called for by the problem tends to throw light. The solution of such problems, it is believed, gives the student experience in working with actual data, the lack of which is often a serious defect in the training of the college graduate who undertakes practical business statistics.

It will be noted that the discussion has been presented in non-technical terms throughout, and no attempt has been made to stress the theory of the subject. While a few indispensable formulae and calculations are given in the Appendix, the more highly mathematical and technical methods have not been discussed. These have their place in statistical thought but fall clearly without the scope of this treatise. Bibliographies given in connection with each subject will be helpful to those interested in more advanced study.

The author is indebted to many persons and organizations for courtesies extended in the preparation of the manuscript. To those who kindly permitted the use of material for illustrative purposes acknowledgment is made at appropriate places in the text.

For valuable criticisms and suggestions relative to the manuscript, the author is deeply grateful to Dr. Roland P. Falkner, Professor of Statistics in the New York University School of Commerce; C. B. Haynes, U. S. Rubber Co.; S. L. Andrew, Chief Statistician, and Dr. R. Belcher and A. H. Richardson, Staff Statisticians of the American Telephone and Telegraph Co.; Martin W. Dodge, Manager of the Industrial Department, Merchants Association of New York; E. E. Lincoln, Chief Statistician, and R. W. Burgess, Staff Statistician of the Western Electric Co.; F. W. Jones, Associate Editor of the *Journal of Commerce;* L. N. Nichols, General Electric Co.; and H. A. Trax, Chief Accountant,

and R. C. Hawkins, Staff Statisticians of the New York Telephone Co.

In connection with the collection and organization of the material acknowledgment is due also to A. G. Pellman, Staff Accountant, N. Y. Telephone Co.; A. M. Shackleton, a former student at New York University; F. P. Hill, formerly President of the Hill-Ware Baking Co.; and Miss M. de J. Cox, Librarian, the American Telephone and Telegraph Co.

The author is especially indebted to H. S. Brown, General Supervisor of Personnel Service of the Western Electric Co., for invaluable assistance during the entire period the book was in preparation; and, finally, to his wife for direct and indirect help and encouragement at all times.

B. F. YOUNG

New York City,
June 10, 1925

CONTENTS

Chapter		Page
I	Introductory	1
	PART I—COLLECTING OR ASSEMBLING BUSINESS FACTS	
II	Facts Required in a Program of Statistics	11
III	Internal Personnel Facts	16
IV	External Personnel Facts	25
V	Internal Financial Facts	33
VI	External Financial Facts	42
VII	Internal Facts of Operations	48
VIII	External Facts of Operations	63
IX	Statistical Sources—General	80
X	Sources for Internal Facts	86
XI	Sources for External Facts	110
XII	Collecting Data—General	130
XIII	Collecting Internal Facts	141
XIV	The Original Collection	149
XV	The Primary Collection—General	152
XVI	The Primary Collection—Use of the Questionnaire	162
XVII	Construction of the Questionnaire	172
XVIII	The Personal Interview	182
XIX	Collecting Facts from Other Business Concerns	197
XX	The Secondary Collection	203
	PART II—ORGANIZING BASIC RECORDS	
XXI	Editing the Initial Primary Basic Records	211
XXII	Types of Basic Records	217
	PART III—PRESENTING STATISTICAL DATA	
XXIII	Presenting Statistical Data—General	231
XXIV	The Textual Method	240
XXV	The Tabular Method	246
XXVI	Constructing a Statistical Table	252
XXVII	The Graphic Method—General Aspects	264
XXVIII	Fundamentals of Graphic Construction	271

Chapters		Page
XXIX	The Bar and Curve	280
XXX	Forms of the Bar and Curve Graph	311
XXXI	The Logarithmic or Ratio Graph	326
XXXII	Maps—Circles—Areas—Volumes—Pictures	338

PART IV—ANALYSIS AND INTERPRETATION OF STATISTICAL DATA

XXXIII	Analysis—General	353
XXXIV	The Method of Sampling	369
XXXV	Frequency Classification	377
XXXVI	Averages—The Mean	402
XXXVII	Averages—The Median	412
XXXVIII	Averages—The Mode	420
XXXIX	Geometric and Harmonic Averages	430
XL	Reliability of the Average—Dispersion and Skewness	436
XLI	Index Numbers	449
XLII	Statistical Relationships—Correlation	464
XLIII	The Time Series	471
XLIV	Nature, Importance and Method of Interpretation	484

PART V—MISCELLANEOUS

XLV	Mechanical Aids in Statistical Work	493
XLVI	Statistics and Accounting	515
XLVII	Organizing for a Program of Statistics	521

APPENDIX

A	Glossary of Statistical Terms	539
B	Suggested Statistical Reports and Statements	549
C	Pertinent Facts Relative to a Selected Group of American Index Numbers	569
D	Logarithms	571
E—1.	Short Method of Computing the Arithmetic Average	577
2.	The Pearsonian Coefficient of Correlation	580
3.	Secular Trend—Method of Least Squares	585
4.	Seasonal Index—Method of Link Relatives	587
5.	Checking Statistical Data	588
6.	Questionnaires	591
F	Problems	593

ILLUSTRATIONS

Personnel, Financial and Operations Records

FIGURE PAGE

1. Employes' Individual Record Changes 92
2. Employes' Payroll Form 93
3. Report of Absent Employes 94
4. Stock Transfer Book 97
5. Dividend Book—Personal Payment 98
6. Sales Journal 100
7. Voucher Record 101
8. Purchase Order 102
9. Purchase Materials Received Report 103
10. Production Order 104
11. Production Schedule 105
12. Idle Equipment Report 106
13. Sales Order 107
14. Shipping Order 108
15. Company Invoice 109
16. Sources of Statistical Data in the Federal Government . . . 116-117
17. Typical Ultimate Primary Basic Record—Long-Hand Form . . . 221
18. The "Tally Method" of Classifying Statistical Data 222
19. The "Check Method" of Classifying Statistical Data 223
20. Separation Record 224
21. A Typical Secondary Basic Record in the Field of Personal Statistics 228
22. A Typical Secondary Basic Record in the Field of Operations Statistics 229

General Statistical Graphs

23. Principal Parts of which the Graphic Chart is Composed 267
24. A Vague Graphic Presentation 267
25. A Clear Presentation of Many Curves 268
26. A Vague Presentation of but Three Curves 269
27. Graduations of Ordinary Foot-Rule Used to Represent Statistical Data 272
28. Equal Distances—Different Values on Different Scales 273
29. The Scale May be Placed Vertically as well as Horizontally . . . 274

FIGURE PAGE

30. Scales May Extend Vertically Upward or Downward or Horizontally to the Left or Right from the Zero Point or Origin . . . 275
31. Division of Chart Field into Quadrants 276
32. Values on the Scales Extend Indefinitely in Both Directions . . . 276
33. Horizontal Scale with Series of Vertical Lines 276
34. Vertical Scale with Series of Horizontal Lines 276
35. Vertical and Horizontal Scales Associated, Each Having a Series of Parallel Lines 278
36. The Simple Bar 280
37. The Component Bar 281
38. Data Sheet 283
39. Typical Application of Scale 286
40. Unequal Space Given to Equal Time Periods on Same Scale . . . 287
41. Good Method of Bringing Out Monthly Figures 287
42. Scale Unit Representing a Space 288
43. Zero Point Omitted 289
44. Figure 43 Redrawn with Zero Line Shown 290
45. Relation Between the Vertical and Horizontal Scales 293
46. Curves too Widely Separated 294
47. Bringing the Curves of Figure 46 into Juxtaposition by the Use of Two Vertical Scales 295
48. Bringing Curves into Juxtaposition by the Use of Two Horizontal Time Scales 296
49. Bringing Curves into Juxtaposition by Means of Two Horizontal Frequency Scales 297
50. Bringing Curves of Figure 46 Together by Breaking the Scales . 298
51. Bringing the Curves of Figure 46 Together by Computing Relatives 298
52. Bringing the Curves of Figure 46 Together by Expressing the Items as Percentages of Yearly Total 299
53. The Rounded Frequency Curve 301
54. Smoothing the Curve by the Free-Hand Method 302
55. The Smoothed Frequency Curve 303
56. Smoothing the Curve by Method of Moving Averages 304
57. Smoothing the Curve by Method of Least Squares 305
58. The Bars of a Bar Graph may be very Narrow 308
59. The Bars of a Bar Graph may be Wide 309
60. Series of Simple Bars 312
61. Series of Component Bars—Bars of Equal Length 313
62. Series of Component Bars—Bars of Unequal Length 313
63. Combination of Series of Bars with the Bars of the Separate Series Grouped Together 314
64. Combination of Series of Bars with the Corresponding Bars of Each Series Grouped Together 315
65. Progress Bar Chart 316

FIGURE PAGE

66. The Gantt Progress Chart 317
67. A Zone Curve 319
68. A Band-Curve—Absolute Numbers 320
69. A Band-Curve—Percentages 321
70. The Zee-Curves 323
71. A Progress Curve Chart 323
72. A Mountain Curve 324
73. A Divergence Chart 324
74. Series Increasing by a Constant Difference Take the Form of a Straight Line on the Arithmetic Grid and a Curve on the Logarithmic Grid 326
75. Series Increasing by a Constant Ratio Take the Form of a Straight Line on the Logarithmic Grid and a Curved Line on the Arithmetic Grid 326
76. A Device for Graduating a Given Line in Accordance with the Scale of Another Line 327
77-79. Logarithmic Graduations from a Plain Scale 327
80. Logarithmic or Ratio Chart Made from Ordinary Ruled Paper by Labeling the Horizontal Lines for Ratios 329
81. Reruling the Grid of Figure 80 to Represent Round Numbers . . 329
82. Use of Two Scales on Logarithmic Grid 331
83-84. Typical Logarithmic Charts 335-336
85-87. Typical Circle Charts 340-342
88. Cartogram—Data Represented by Cross-Hatching 343
89. Cartogram—Data Represented by Size of Dots 343
90. Cartogram—Data Represented by Frequency of Dots 344
91-97. Typical Picture Graphs 346-349
98. Combination Graph—Map and Bars 350
99. Combination Graph—Curves and Columns 351
100. Personnel Facts to be Associated 361

Frequency Graphs

101. The Frequency Curve 379
102. A Continuous Frequency Series 380
103. A Discontinuous Frequency Series 382
104. The J-Shaped Distribution—Personnel Data 394
105. The J-Shaped Distribution—Operations Data 395
106. Comparison of Frequency Series—Absolute Form 396
107. Comparison of Frequency Series—Percentage Form 397
108. Cumulative Frequency Curves or Ogives 399
109. A Less-Than Cumulative Curve 419
110. A Tri-Modal Distribution 425

INDEX, CORRELATION AND TIME SERIES

FIGURE PAGE
111. Interpretation of Index Series 462
112. Correlation—Curves that Parallel 467
113. Correlation—Curves that Mirror 468
114. Trend Line in a Time Series by the Methods of Semi-Averages and Moving Averages 474
115. Trend Line in a Time Series by the Method of Least Squares . . 477

MECHANICAL AIDS

116. Tabulating Card 494
117. Dual Tabulating Card 495
118. Method of Representing Data on the Tabulating Card 497
119. Electric Key Punch 498
120. Electric Key Punch with Duplicating Attachment 498
121. Verifying Key Punch 499
122. Electric Sorting Machine 500
123. Electric Accounting Machine 500
124-125. Scales Illustrating the Use of the Slide Rule 506
126. The Slide Rule 507
127. Multiplication Process on Slide Rule 511

ORGANIZATION CHARTS

128. Organization Chart—the Centralized Type of Statistical Organization 530
129. Organization Chart—the Decentralized Type of Statistical Organization 531

STATISTICAL REPORTS AND STATEMENTS

130-131. Typical Secondary Basic Records 550-551
132-136. Typical Statistical Reports in the Personnel Field 552-556
137-143. Typical Statistical Reports in the Financial Field . . . 557-562
144-147. Typical Statistical Reports in the Operations Field—Purchasing, Producing, Marketing 563-566

STATISTICS AS APPLIED IN BUSINESS

CHAPTER I

INTRODUCTORY

Rapid Development of Statistics.—The subject of statistics as a special study is of comparatively recent origin. It has been only within the last generation or so that our institutions of higher instruction have offered systematic courses in the general subject of statistics and that books have been published, devoted entirely to the technical subject. But the growth has been remarkably rapid. In the colleges and universities of the country today we find extensive statistical departments and courses which aim to discuss the subject in full and complete form. The development of the subject beyond the walls of the colleges and universities has been even more marked. In business and commerce, in educational and religious institutions, in civic, social, economic, and governmental activities, statistics has found an extensive and, we believe, a permanent place.

Statistics and Business Management.—In recent years the management and direction of a business organization has been the subject of much study and investigation. Efforts to improve business methods and increase the efficiency of the organization have been made along many lines. Books have been published on various phases of management, magazines have been filled with articles outlining the experiences of some successful manager or portraying particular aspects of the management of a business concern, and periodicals have grown up whose principal purpose is the presentation of literature on business management. Likewise, organizations have been formed under various names whose sole aim is the consideration of problems of management. We are

resorting to all sorts of devices today to increase the efficiency of management in business.

Along with this development has come an appreciation of the service of statistics in management and a great deal of attention has been given to the application of statistical methods to practical business problems. This book is written primarily to portray in systematic and orderly manner just how the elementary principles of statistical methods may be serviceable in the management of a modern business organization. In other words, statistics, as of interest in these discussions, is essentially a body of business facts and the methods for handling them by means of which we aim to extend our control in the field of business management.

Scientific Tendency of Modern Industrial Management.—Modern industrial methods have tended to make business a science. The scientific method has been carried to a greater or less extent into practically every phase of business activity and more and more each day is it being extended to a wider field of problems in the realm of practical affairs. The former method of directing the affairs of a business enterprise—fashioning its policies, solving its important problems, developing its plans for the future—on the basis of "long-headedness" or "business intuition," is fast passing, if indeed it has not already passed in the larger industrial enterprises. The method of "guessing" has proved inadequate and has given way to the more accurate, the more exact scientific method. In progressive industrial organizations the scientific method, imperfect as it may be in its application, is depended on to guide the business enterprise through the many intricate and complicated difficulties of modern industrial life to success.

Many things have contributed to the need for this method. The keener competition of modern business has compelled men in general to operate on a smaller margin of profit. In the production of his goods or the operation of his business the business man has to reckon with a laboring class and a consuming public much more enlightened than he has been accustomed to deal with in the past. Labor is constantly demanding and organizing to obtain greater consideration from the business organization in the form of higher wages and better employment conditions. This fact adds ma-

terially to the cost of production. The consuming public, moreover, is keen to prevail against a retail price which seems unjust. Through public opinion expressed partly in legislative action, as in the case of governmental regulation of public utilities or the functioning of the Federal Trade Commission, and partly through non-governmental means, as expressed in the public press, or at mass-meetings, the producer is forced to confine his retail prices within certain rather vague limits, keenly sensed, however, by the consuming public, beyond which the price is unjust in the eyes of the consumer. The consuming public will no longer pay, without protest, whatever prices they may be asked to pay.

On the one side, therefore, is the laboring class making demands which increase the cost of production. On the other is a consuming public keenly sensitive to possible injustice at the hands of the producer, constantly agitating to keep retail prices down. The producer finds himself between the upper and nether millstone. His only way out is improved methods of production. He must find some means whereby he can meet the added cost of production required by the demands of labor without shifting this additional burden to the consuming public in the form of higher retail prices. In other words he must increase the output per unit of energy expended on the one hand or reduce the cost of marketing on the other, to the end that the per unit cost remains relatively constant even though the labor cost has increased substantially. To realize economies through improved methods of management the producer employs the scientific method.

Furthermore, the enormous growth of business enterprises in the last few years has contributed in no small way to the need of more efficient methods of management. In many instances the business organization is so large today that it cannot be personally supervised by the management as was the case but a few years ago. The assets of business organizations have changed from hundreds of thousands to hundreds of millions of dollars and the income and expenditures from thousands to millions. The number of employees has shifted from a few hundred, who were known personally to the manager, to hundreds of thousands, a large number of whom may not even know the name of the president or general manager, and practically none of whom are known personally

to the managing officers. Many business organizations of today are operated on such a large scale that the management cannot ascertain the facts and essentials of the concern by the methods of personal inspection which was followed in former years. The management must get acquainted with the business in other ways and must employ other methods to direct its affairs.

To recapitulate, the scientific method is the hope of modern business. On its use depends the success of the enterprise. Without the safe and sane management which it makes possible, no business concern can long endure the strain and stress of modern industrial life. Even though its application is imperfect in comparison with its use in the theoretical laboratory, nevertheless the essential elements of the method are followed carefully and important decisions are made only after careful consideration of the facts scientifically organized and interpreted.

Availability of the Facts.—Whatever else may be said of the scientific method, it is obvious that a first requisite of its use is the availability of the facts. Consequently, coincident with the development of the use of the scientific method in business, provision must be made to have the detailed facts of the business readily available. Complete and comprehensive data covering every phase of the business—combined, summarized, associated, compared, and otherwise organized to reveal their bearing on the important problems of the enterprise—must be at hand. Rates of pay, for example, are important facts in themselves but their use is limited unless they are organized according to departments, divisions, and occupational classes; summarized in totals and subtotals; expressed in averages, percentages, or ratios for purposes of comparison; or associated with other groups of facts such as periods of service, age, labor losses, and so forth. Such an organization of data comprehends not only the essential detailed facts in connection with each of the fundamental phases of the individual establishment itself but embodies all external information related in an important way to the management of the business.

Nature and Function of a Statistical Organization.—Recognition of the importance of the facts in business management

as noted above—emphasized no doubt by the extensive use of statistics in the management of the recent war organization—has led to the establishment of organizational facilities for rendering this service on an adequate scale. Whereas heretofore information had been culled unsystematically, as special needs developed, from nooks and corners, wherever persons or records could be found that would give something in the neighborhood of the facts required, at present a thoroughly developed organization, well-manned and well-equipped, is provided to record these facts systematically and to have them ready in proper form for the services of those whose activities require them. The assembling and organization of detailed facts in regard to number of employes, absences, latenesses, labor losses, wages, purchases, sales, deliveries, and other phases of the business is set aside as a distinct function and a separate organization is provided for accomplishing it.

In other words, the development of systematic information required in the scientific management of the business is a special and distinct phase of the operation of a business organization, just as purchasing, producing, and marketing are distinct phases. As such, it requires a separate and distinct organization specially manned with technically trained employes and equipped with adequate mechanical aids to accomplish the work. The function of the statistical organization is, broadly, to maintain systematic information in regard to the fundamental processes or activities of the business and to furnish this information in proper form to those in the organization who have need for it. Specifically, these functions will be as follows:

1. To collect and record statistical data which are required in the management of the business.
2. To analyze these data.
3. To present these data in proper form and distribute them to the individuals concerned.
4. To aid in the interpretation of these data.

The statistical organization is therefore a producing organization. Its raw material is the facts of the business, its processes are analytical, and its finished product completed statistical studies designed to serve the management in their effort to direct the

affairs of the business scientifically. The importance of this service to modern business management can scarcely be over-emphasized. More and more, in fact, the success or failure of the modern business enterprise depends on the efficiency of this service.

Nature and Meaning of Statistics.—A great variety of definitions have been offered. These range from the narrower conceptions which associate the subject with a definite body of facts and insist on *statistics of something,* or require that statistics comprehend only facts relating to the state, or confine the term to the process of counting or to a particular analytical statistical device; to the more scientific conception which defines statistics as "aggregates of facts, affected to a marked extent by a multiplicity of causes, numerically stated, enumerated, or estimated according to reasonable standards of accuracy, collected in a systematic manner for a predetermined purpose, and placed in relation to each other."

Without any attempt to offer a technically worded definition, we shall adopt the practically serviceable conception and use the term statistics to indicate *numerical data together with the body of more or less specialized and standardized methods required in the collection, analysis, and interpretation of these data.* In other words, in the practical field statistical data are but the detailed numerical facts of the business and statistical methods the body of rules and principles governing their analysis and interpretation —the means whereby we determine the bearing of these facts on the important problems of the business. In this sense we may use the term to indicate either the data, the method, or both data and method.

In the past the character of practical statistical work coincided with the first concept noted above. It consisted primarily of collecting statistical data and filing them away with little or no attempt to utilize them in practical affairs. Such work consisted of little more than the bare collection and simple classification of data and the statistical organization little more than a "figure factory." In this way statistics came to be known as a very uninteresting, useless, and dry subject with little or no practical interest or significance.

But the modern conception of the term views statistics practically as a means to an end. Today we accumulate statistical data, and analyze and interpret them, simply as a means of solving important problems. This conception has transformed the subject into a very live and fascinating one. In a lecture before the Harvard School of Business in 1913 on the subject "The Functions of Statistics in the Telephone Business," Walter S. Gifford, now President of the American Telephone and Telegraph Company, said: "In so far as all science and scientific search for truth can be said to be boresome and apart from human interest, just so far and no farther can Statistics be so classed."

Statistical data are but the facts of the business and as such are indispensable to scientific business management and to the "scientific search for truth" in the realm of practical affairs. If statistics is confined to the mere handling of figures, it is indeed a dry subject; but to the man who is mentally alive and who looks on statistics as an effective tool in the solution of fundamentally live and important problems, statistics loses its dryness and becomes one of the livest and most fascinating of subjects.

It should be noted in this connection that the fundamental basis of all statistical developments is the process of counting. Numerical facts obtained through this process constitute the raw material of statistics. Hence, in the use of the word "facts" in these discussions we shall have in mind statistical or numerical facts.

The Use of Statistics in Business.—The business executive is no longer compelled to make a blind guess as to what his production should be, or what his sales will probably amount to, or whether or not a certain advertising campaign has been worth while, etc. He can guide his judgment in these matters by carefully worked out statistical calculations which will enable him to make a much closer guess in regard to these and many other equally important problems. In short, the policies of business organizations of the country today are determined on the basis of facts. Some illustrations will serve to make the picture more concrete.

The first consists of a problem which was reported as having been assigned to the statistical staff of one of the large business

offices in New York City. This company does an extensive business in both the United States and foreign countries. Some few years ago the statistician was called into the office of the vice-president and general sales manager and assigned this problem: "What are the business possibilities in the United States of the kind of business in which we are engaged?" In other words, the firm wanted to know broadly what the extent of their business would be in the United States if they had no competitors and the country were developed to a maximum point in their line of business.

In undertaking the solution of this problem the statistician set up three main inquiries—"How much business is our firm doing?" "How much business are our competitors doing?" and "How much undeveloped business is there in the country?" The first of these problems was easy; the second was a little more complicated; and the third called for a very elaborate statistical investigation, involving a study of the natural resources, the development of communication and transportation facilities, the probable change in the composition and the number of population, and requiring some three years to complete. This inquiry was then supplemented by the attempt to determine what portion of this potential business the company could reasonably expect for itself.

This problem is typical of the more elaborate research studies in the practical field. Elaborate investigations of this kind are confined for the most part to the larger business concerns but investigations similar in character are undertaken freely by smaller establishments. Even the retail grocer may profitably inquire what the maximum consumption of groceries is in his territory and what portion he should reasonably expect.

Problems of a little different nature are illustrated by statistical developments designed to keep before the executive a clear picture of the conditions of the business. Such a statement will generally consist of three parts representing, respectively, (1) the present status of the business showing for stocks, sales, and deliveries, the actual facts compared in absolute amounts and percentages with the requirements; (2) a future estimate showing a similar comparison at some future date, say three to six months; and (3) a comparison of current production with that required to meet de-

liveries. Figure 146 given in Appendix B shows the details of such a statement.

These illustrations will suggest in a very broad and general way the part that statistics plays in the management of the modern business concern. Probably no one is in such urgent need of statistical data and devices for handling them as the man charged with the actual operation of a business organization. He has to meet a great variety of problems and often in a very short period of time. He may have but a few data and have to base a decision on them or he may have a large group of data and have to reduce them to a condensed, representative form which the mind can grasp. In any event he turns to statistical data for such assistance as he may derive from them.

It should be emphasized in this connection that the need for a knowledge of statistics is not confined to specialists. In most concerns the statistical work is scattered throughout the organization and all the supervisory and executive forces who have to do with the compilation of statistics or with their interpretation in deciding practical problems should have a working knowledge of elementary statistical methods, regardless of what department they may be located in.

The extensive use of statistics in business has created a demand for persons trained in the elements of the subject, and one very practical reason for offering courses on the subject in our colleges and universities is the need for training such persons. There is no inclination to suggest that the statistical method can be used in the solution of all the problems with which the business man is confronted nor that it is the only method. The field to which it can be applied, however, is growing continuously wider, and it will undoubtedly continue to expand.

Part I—Collecting or Assembling Business Facts

CHAPTER II

FACTS REQUIRED IN A PROGRAM OF STATISTICS

Detailed Facts the Foundation of a Statistical Service.—The foundation of a statistical service in a modern business concern is the detailed facts of the business. They are the raw material out of which such a service is built up. If the service is to be adequate, the foundation must be complete and comprehensive. The facts available must, therefore, cover every phase of the business. Facts in one field do not suffice for other fields. The most thoroughgoing organization of labor facts, for example, does not furnish data to solve problems in the financial field. Facts on rates of pay and labor turnover will not be pertinent to a consideration of the question as to whether new financing should take the form of a stock flotation, a bond issue, or a bank loan. Likewise, financial and operating facts will not suffice in the solution of personnel problems. Each field of problems in business management calls for a distinct group of facts, all of which must be available if the statistical service is to be adequate and efficient.

Need for Statistical Information.—The need for statistical data is evidenced at every turn in the management of a business enterprise. Important problems are constantly before the executive, the answers to which are based on systematically developed statistical information.

What labor force is required properly to conduct the business? How are the employes distributed as to age, sex, occupational class, period of service, and so forth? Is the company getting from them the service that it is paying for? Are the employes being properly paid? Are they generally in good physical condition?

In another field, what is the source of capital invested in the business? What does it cost? How is it used? What return is realized on it?

Likewise, what is the quantity, rate, and cost of production? How does the rate of output compare with the requirements? What is the quantity, selling price, and cost of goods sold? What profits have been realized according to classes of goods, salesmen, and territories.

These and many other problems of equal importance require constant attention. The basic record data outlined in the following chapters are designed to furnish detailed facts on the basis of which these problems may be properly disposed of.

Fundamental Classification of Business Facts.—From a broad point of view, the fundamental elements which are generally applicable to all business organizations are (1) the personnel, (2) the finances, and (3) the operations. Each of these fields requires a specific group of facts, which may be used alone or in association with facts from other groups. The six chapters immediately following offer a more or less comprehensive group of the classes of personnel, financial, and operations facts which will be required as a foundation for a statistical service in business concerns generally.

More specifically, the general classification of the detailed facts in these chapters is as follows:

The Personnel—	The working force
	The wages
	The employment conditions
The Finances—	The acquisition of capital
	The use of capital
	The return of capital
The Operations—	Purchasing
	Producing
	Marketing

The fundamental classification is worked up separately for internal and external facts. Supplementing the fundamental grouping, special consideration is given to the development of these basic data in historical studies, future estimates, and stand-

ard or normal facts. Furthermore, certain additional facts are noted which are not included in the outline.

Historical Facts.—The facts contained in the systematic records may be worked up in historical form for the period covered by existing records. It will be advisable at times, however, to transfer this material from the basic records to records whose primary outline is historical.

In addition to the facts contained in the primary basic records, other classes of data will have an historical interest and these will generally be provided for in historical form, taken as derived or computed figures from primary or secondary sources.

Future Estimates.—Future estimates are considered in two broad classes—current or working estimates and long-time estimates. Those in the former group partake of the nature of definite plans for the immediate future and are developed in fairly great detail. Those in the latter group usually take the form of long-time trends and are confined to a relatively few facts of basic importance.

Current or working estimates generally cover only a year or two in the future while long-time estimates extend to five, ten or fifteen years or a longer period depending on the character of the business.

Standard or Normal Facts.—This group of facts is offered for a two-fold purpose—to emphasize the importance of establishing standards against which the actual facts of the business may be measured, and to indicate by means of illustrations the kinds of standard facts that should be developed in each general class of data.

Fundamentally, standard facts are designed to reflect the status of the business when it is as we should like it or should expect it to be. On the one hand, they are indicative of a sort of ideal, and, on the other, they represent the expected status of the business as determined by a scientific evaluation of the facts. In any case, standard or normal facts are highly important in business management and their use cannot be too strongly urged.

Additional Facts.—Included in this group are additional classes of data which, because of their relative unimportance for business concerns generally, are omitted from the regular outline. In certain companies all or part of them may be currently recorded on the basic records as being somewhat extensively required.

Offered as a Type Only.—It will be kept in mind during this discussion that any list of detailed facts purporting to typify the facts of all kinds of business organizations must be general in character. The lists given in the chapters immediately following will not precisely fit any specific establishment. They are intentionally made rather comprehensive. The primary aim is to build up in systematic form a group of facts which may be helpful by way of suggestion to those who undertake the development of a program of statistics in any business organization.

The limitations of the group here outlined, in its literal application, are obvious. In some concerns large sections of the group may be inapplicable; for example, the facts on stock and bond flotations would not apply to a non-corporate business. In some other concerns perhaps entirely new groups of data would have to be worked up. Specialized enterprises such as the motion picture industry, fertilizer production, warehouse service, and the retail tobacco business will require the addition of other groups of facts. Likewise, small establishments will generally not require such a comprehensive group. To develop a variety of lists of facts suited more specifically to different types of business enterprises would present a problem too far-reaching in scope and too complicated in character to be included in a treatise of this kind. The detailed facts listed in the several classes may vary but the broad classification—personnel, finances, and operations—is generally applicable.

We should ever bear in mind that the assembling and organization of statistical data in a practical business concern can be readily overdone. It cannot be too strongly emphasized that in practical statistical work the development of statistical information is never an end in itself; it is wholly a means to aid in the solution of practical problems of management. In order to maintain this conception, every group of facts that are considered for entrance on the

basic statistical records should meet the final test of service and in no case should the statistical work of the company be permitted to grow beyond the point where the cost of the statistical service exceeds the benefit derived from it.

Furthermore, it is important to note that actual programs of statistics in business concerns will show many variations from the type suggested herein both as to form and character of data. Frequently, the statistical service "just grows" and when this is the case its basic structure will take a variety of forms. The statistical service may not be centralized either as to organization or as to supervision, or, indeed, it may not even be systematized. Sometimes the statistical work of a company develops at different points in the organization concurrently and there is no effort to systematize the information or to establish a comprehensive program of facts. The effort to improve such situations should be made with full appreciation of the practical factors involved.

Small and Medium-Sized Companies.—Particularly will the program outlined herein have to be modified to meet the needs of the small and medium-sized company. While the fundamental classification of facts will be equally pertinent for all sizes of concerns, the body of detail will be much less in programs of statistics for the smaller organizations. In such cases, certain classes of facts will be selected from the outline on the basis of which the statistical service will be developed.

The position is taken throughout this treatise that statistics are of tremendous importance in the management of medium-sized and small companies as well as large corporations. Even the small retailer is confronted by many important questions: How much business, actual and potential, is there in the territory which the store claims to serve? What portion of the total should the store get? What portion does it actually get? What are standard delivery costs, wastage, revenue, profit, and so forth? Statistical data will be serviceable to this simplest type of business organization in connection with the solutions to these and many other equally important questions. Likewise, the small manufacturing company will find many occasions to use statistical data.

(For references see page 79.)

CHAPTER III

INTERNAL PERSONNEL FACTS

Basic Record Facts Only.—The facts in this group will be organized under three main heads—the working force, the wages, and the employment conditions: in other words, the employes, the moneys paid to them, and the conditions under which they work.

The detailed facts that would be required in the proper handling of problems arising in connection with each of these three phases of the personnel of a business organization are listed below. This list includes only the detailed facts that should be entered on the basic records and does not include important facts which may be derived from them.

The several items related to each main heading are noted individually. The added significance of these facts when associated with each other with other facts of the business is discussed in Chapter XXXIII.

I. THE WORKING FORCE

A. *Facts relative to the Number and Characteristics of the Employes.*

Basic records should show the following information with respect to each individual employe:

1. Regular, part-time, or temporary employe
2. Day or night worker
3. Office or factory worker
4. Married or single
5. Health class—*i.e.*, the class designated by the medical department on initial examination at the time of the employe's entry into the service or at a later time
6. Sex
7. Date of birth
8. Date of entry into the service of the company
9. Occupational class
10. Nationality

11. Education—General— Grade school
High school
Special school (kind specified)
College—name of institution, dates of attendance or graduation
Degree
Technical—Name of institution, dates of attendance or graduation.
Character of training
Degree

12. Experience—Employment in the company prior to time of present engagement
Previous employment in other companies
Trade skill

This material should be recorded on the primary basic records of the statistical organization in such form as to permit the enumeration of the facts or segregation of the data weekly, monthly or quarterly and by departments and divisions or by other time periods or organizational units.

This list might be extended by the addition of facts which are less generally important such as the religion, number of dependents, and state of citizenship of the employe; his ability to read, write and speak the English language; date of eligibility to the various provident plans of the company; and so forth.

B. *Attendance Facts*

The primary basic records for attendance facts should be organized to show both absences and latenesses associated with such of the items listed under I-A, "The Number and Characteristics of the Employes," and such of the wage facts listed as may be desired. For example, in the development of the basic records, it would be generally advisable to provide for showing absences and latenesses separately for male and female employes, day and night workers, office and factory workers, and the various health classes. In addition, however, to such distribution among these items as may be chosen, the following facts should be available separately for each individual absence or lateness:

Length
Cause

Deduction from pay
Date
With or without previous permission

Furthermore, certain general facts should be recorded such as the following:

Attendance bonus—Money payments
Extra vacation allowance
Leave of absence
Increase in pay

Penalties for poor attendance—
Decrease in pay
Decrease in vacation
Retardation in pay advancement

C. *Labor Turnover Facts*

The primary basic records which contain the raw material bearing on labor turnover should be organized to permit the working up of both additions to and losses from the force in association with such of the classifications listed under I-A, "The Number and Characteristics of the Employes," and such of the wage facts listed under II as may fill the requirements of the particular situation.

They should also provide the following additional facts separately for each individual employe:

1. Additions—Date of employment
Source of employe
Methods of obtaining employe (*i.e.*, employment agencies, newspaper advertising, through employes, etc.)
To fill vacancy
To fill a new position

2. Losses— Date of leaving
Cause
General classification (*i.e.*, dismissal, resignation, release, etc.)
Destination if going to other employment

3. Transfers—Date
Present location, department and division
New location, department and division
Reason (*i.e.*, promotion, find work better suited to employe, etc.)

Present salary
New salary

4. Cost Data—The data required in estimating the cost of labor turnover will be assembled through special studies.

II. THE WAGES

Basic records should contain the following information with respect to each individual employe:

A. *Rates of Pay* (authorized base rate)—
- Flat rate
- Piece rate
- Increase— Date
  - Amount
  - Classification (*i.e.*, automatic, general, individual)
- Decrease—Date
  - Amount
  - Classification (*i.e.*, general or individual)

B. *Overtime Payments*—Date
- Amount of payment
- Amount of time
- Classification (*i.e.*, extra hours, Saturday or other half-holiday, holiday, Sunday)
- Basis of payment (*i.e.*, straight time, time and one-half, double time, other)
- Form of payment (*i.e.*, money, time off, extra vacation allowance)
- Nature of work

C. *Additional Compensation*

1. Bonus— Date
   - Amount
   - Classification (*i.e.*, general, applying to entire group, or special, such as attendance, efficiency or anniversary)
2. Commission—Date of payment and period covered
   - Amount
   - Rate

3. Profit-sharing Payment—
   Date
   Amount
   Basis of payment
4. Differentials—Date
   Amount
   Classification (*i.e.*, for Sunday work, Saturday or other half-holiday, holidays, night, other)

D. *Deductions*—Date
   Amount
   Classification on basis of purpose or cause (i.e., time lost, penalty, fine, dues, overpayment, payment on company stocks or bonds, etc.)

E. *Total Payroll*

The facts on wages outlined above should be so recorded as to permit organization on a daily, weekly or monthly basis, and by departments, divisions, or in such other form as may be quired. They should also be so organized as to permit association with the items listed under I, "The Working Force."

III. EMPLOYMENT CONDITIONS

A. *Hours of Work* —Number of hours in regular business day and week
   Length of lunch period
   Length of relief periods
   Time for beginning and ending the business day

B. *Vacation Practice*—Regular allowance
   Special allowance recorded for each individual employe (such as that given as bonus for attendance, period of service, or efficiency)
   Additional employes required on account of vacation, showing the number and the amount of pay

C. *Employes' Service Activities*

The statistical records should contain detailed information in regard to the various forms of employes' service activities particularly those which are sponsored by the company. In connection with the *Employes' Benefit Organization* whose

general purpose is to provide certain types of systematic benefits such as those given in connection with sickness and accident disability, death, and old age, typical of the more important facts are the amount of moneys disbursed; class of case; nature and length of illness; and nature, cause, and severity of accidents. These facts should be recorded so as to permit the development of the data in association with such other personnel facts listed herein as sex, age, occupational class, health class, hours on duty, period of service, and rate of pay.

Detailed facts as required should be provided in connection with other employes' service activities such as Medical Service, Savings and Loan Plan, Co-operative Store (including sale of discounted company goods), Restaurant Service, and Employes' Stock Plan.

The facts listed under *attendance, labor turnover, wages, and employment conditions* should be so recorded as to permit the development of the data in association with pertinent facts listed under I-A, "The Number and Characteristics of the Employes." Provision should also be made for the association of wages with facts on attendance, labor turnover, and certain groups of facts under employment conditions. More detailed discussion of the association of various kinds of facts is given in Chapter XXXIII.

Cost Data in the Field of Internal Personnel Statistics.—The actual money cost to the company of absences, latenesses, labor turnover, vacations, holidays, employes' service activities, and so forth, can generally be obtained directly from the data listed above under these headings or by the association of these data with wage facts. In certain cases, special methods will need to be developed for working out cost facts. The data required in estimating the cost of labor turnover, absences, and employes' service activities, for example, will generally be assembled through special studies.

Whenever the required costs cannot be obtained from the data listed above under these headings, taken either alone or in association with other data listed in this chapter, they should generally be developed by special study rather than by an attempt to maintain the required data on the current records. In certain

cases it may be advisable to work up basic costs per unit which can be applied to current facts to obtain the required costs.

Historical Data.—Historical data in the field covered by the outline and for a period prior to the establishment of the detailed records as indicated should usually be worked up through special studies. In case an inquiry should call for the development of certain data from past records over a period of time, it is advisable to make a rather complete transcription of the type of data required and record it systematically, even though some of the facts are not needed in connection with the particular inquiry. In collecting data, for example, relative to an inquiry which calls for only the total number of employes by months over the past ten years, it is generally worth while to transcribe figures showing not only the total number of employes but the total distributed by sex, occupational class and departments or divisions of the organization, while the records are at hand. This, of course, applies only to periods of time for which the detailed and systematic records outlined herein have not been developed.

Furthermore, certain important data may well be worked up as time permits even though no special inquiry requiring them has arisen; such as the number of employes, absences, latenesses, labor turnover, rates of pay, overtime rate, additional compensation, and sickness disability cases, all of which should be distributed by sex, occupational classes, office and factory workers, or other classes.

Future Estimates.—Future estimates of internal personnel facts will be of two general classes, current or working estimates and long-time estimates. The former group will contain estimated personnel facts relative to the immediate future—the next month, quarter, or year—and will take the nature of more or less definite plans for the future. The latter group will contain estimated personnel facts for the relatively remote future—the next 5 or 10 years or even a longer period—and partakes more of the nature of the future trend or tendency, the results being much less definite than those in the former group.

The current or working group will include such detail as esti-

mates of the number of additions to and losses from the force (distributed by sex, occupational class, or other detail), rates of pay, the cost of employes' service activities, and so forth.

In the long-time group, future estimates will be confined generally to the development of future wage trends and the number of employes the company will require in the future. Along with the future trend of wages in general, the statistical department should develop such future estimates of specific rates of pay of the individual establishment as may be needed. Future estimates of the number of employes the company will require will be developed separately for broad occupational groups.

Standard or Normal Facts.—There are certain standard or normal facts which should be developed in the field of personnel statistics. They are generally of the computed or derived form and are mostly developed on the basis of external data. They are noted here, however, because of their close relation to and importance in connection with internal facts of the business. At least the following standard or normal facts in the field of personnel statistics should be available from the statistical records:

1. Normal or Standard Labor Turnover.—A normal or standard labor turnover figure should be developed, perhaps preferably in percentage form. Such a figure might be kept at hand for the particular individual establishment and for the industrial class as a whole to which the individual company belongs. Generally, the former will better serve the needs of the company, and the latter will enter into the determination of it. Further requirements in this field will lead to corresponding standards separately for additions to and deductions from the force. Additions and deductions due to creating a new department or eliminating an old one would of course not be included in these standards.

2. Normal or Standard Absence.—This standard is preferably expressed in terms of the average number of days per employe, distributed between men and women employes, and by any other groups that may be desired such as occupational classes, health classes, nationality, married and single employes. Such standards might also be developed for absences due to specific

causes. Actuarial figures available from insurance companies will be very helpful in establishing standards for certain specific causes such as sickness and accident.

3. Normal or Standard Wage.—A normal or standard wage should be currently available against which the actual wages of the company can be measured. Standards should be provided separately for such different groups of employes as may be considered worth while, particularly by occupational classes and sex groups.

4. Additional Standards.—The foregoing standards are noted as illustrative of this type of data but are not presented as a comprehensive list. The need for additional standards will be apparent as the program of internal personnel statistics is developed in the particular case.

Additional Facts—General Information.—In addition to the systematic list of facts outlined above there are many data that might be included, such as those relating to the employes' representation plan; the reduction of number of employes due to the introduction of labor-saving machinery and other improvements in methods; comparative efficiency of individual employes and groups of employes; methods and frequency of wage payments, including legal requirements pertaining thereto; and real wages of the employes of the company according to some fixed standard.

It is believed that in general the service to be derived from such data does not warrant their development and maintenance on detailed basic records. Should the need justify it in any particular case, the facts could be readily listed after the fashion of the outline given above and provision made for their proper recording.

(For references see page 79.)

CHAPTER IV

EXTERNAL PERSONNEL FACTS

The Company and the Business Community.—The individual business establishment will find it desirable from time to time to make comparisons of its internal personnel facts and conditions with corresponding facts and conditions within the business community. Its personnel problems are broadly the same problems which other businesses of the locality have to face, and its personnel practices are generally determined by theirs. The individual company cannot depart greatly from the practices and policies of the community in which it is located without suffering materially thereby. It is highly essential, therefore, that the management of the individual concern know where it stands with respect to the fundamental elements of its personnel practices and policies. This is particularly true with regard to certain fundamentals such as rates of pay, hours of work, vacation, and employes' service activities.

Such information can be arrived at only when external personnel facts are available for comparison with those within the company. It therefore falls to the function of the statistical organization to provide adequate service in the field of external personnel facts. For the business community as a whole, particularly with reference to business establishments of the same general industrial class, data on (1) the working force, as suggested at the beginning of Chapter III, (2) their wages, and (3) the conditions under which they work, should be available in a form which will permit comparison with the corresponding internal facts. This will hold equally true for facts related to the current situation, past history or future trend, although, generally, facts in the last two fields can be adequately handled through special developments.

Groups Considered, Not Individuals.—The outline followed in the organization of personnel statistics within the company will

serve as a guide in the collection of external facts. To obtain information in such detail, however, would probably be neither possible nor desirable. In the internal field the unit of consideration is the individual employe; in the external, a group of employes. Whereas within the company our program provides for the current recording of certain personnel facts such as rates of pay, absences, and latenesses regarding the individual employe, the statistician's interest in the external field is usually confined to broad, general considerations and the basic data are generally those related to groups of employes. These groups may be definite, such as the employes of a specific outside business concern, or indefinite, such as those represented by newspaper advertisements or the records of employment agencies. In connection with the latter the facts will sometimes refer to individual employes. The outline below has been developed primarily with the definite group in mind.

Less Comprehensive in Scope Than Internal Facts.—The facts listed here as being desirable for basic records will be less comprehensive in scope than those contained in the list of internal facts. A greater portion of external personnel facts should be handled through special studies rather than through the medium of current records. For example, facts relating to wage deductions, medical supervision, savings and loan plan, cooperative store, and employes' stock plan, should be assembled through special surveys and hence are omitted from the list given below. The facts contained in the list should be systematically collected and recorded as a regular procedure.

Derived Facts as Basic Data.—Furthermore, a more extensive use can be made of derived or computed figures in the external field, particularly in the case of historical data. The list of basic record facts given below is therefore divided into two separate types of data—(1) primary basic facts related to groups of employes, and (2) derived or computed figures.

Form of External Basic Data.—The form of external personnel data will be somewhat different from that of corresponding

internal facts. The external facts will generally exist in a partially summarized form instead of appearing for each individual employe, although facts related to indefinite groups such as those represented by newspaper advertisements or the records of employment agencies may be an exception. For example, the records containing the personnel facts for a specific business concern in the community will not generally show whether a particular employe is regular, part-time, or temporary, but rather will they indicate that, of a certain total number of employes, so many are regular, so many part-time, and so many temporary. The same is true for the other items listed. The general form is merely a total number distributed by certain specific groups.

Complete Data Not Always Available.—It should be noted also that comprehensive external data cannot always be had. Many concerns do not have the facts in such detail on their records and others are reluctant to part with them although the number in the latter group is rapidly growing smaller. This outline of facts represents the completed program toward which we should strive; and in most communities, particularly among the larger concerns, a well-directed, conscientious effort to obtain these facts will produce surprisingly satisfactory results.

Primary External Data.—Facts in this general group will usually be detailed basic facts which have not been finally summarized or analyzed but which have to be worked up before they can be used for the purpose intended. They are in contrast to computed or derived data which are discussed below.

The details required will be generally similar to those listed in Chapter III above. The records related to definite groups will usually show so many employes in each of the classes listed in Chapter III. For indefinite groups such as those represented by newspaper advertisements and records of employment agencies the data may appear in the same form, particularly those from the agency, or the records may carry the detail which is to be worked up into this form. Provision should be made in the collection and recording of these facts for such association among the several groups of facts as may be desired.

Computed or Derived Facts.—Computed or derived data in the external field are frequently serviceable and the statistical organization should provide for basic records containing such of them as may be pertinent. A great variety of such data will be available from literature and from the records of other concerns. Definite and systematic efforts should be made currently to assemble and record certain of these facts while others can be satisfactorily handled through special studies as the need for them arises. The list given below is suggestive of the type of computed or derived facts which should be available at all times from the records of the statistical organization. The facts listed should be available for individual representative business establishments or industrial groups in the community or for the community as a whole.

1. Attendance Facts
   - Average length of absences.
   - Average length of absence per employe.
   - Average number of absences per employe.
   - Average number of latenesses per employe.
2. Labor Turnover Facts
   - Actual percentage labor turnover.
   - Standard or normal percentage labor turnover for the several industrial groups.
3. Wage Facts
   - Average authorized or base rate of pay.
   - Average earnings.
   - Percentage increase in base rates or earnings over a specified period of time.
   - Overtime payments.
     - Average payment per employe working overtime.
     - Average payment per employe in company.
     - Average amount of time per employe working overtime.
     - Average amount of time per employe in company.
   - Additional Compensation (bonus, commissions, profit sharing, etc.).
     - Average per employe.
4. Employes' Service Activities
   - Average length of disability cases receiving benefits.
   - Average payment per case and per employe eligible to benefits.

Actuarial figures showing the expected sickness and accident disability per employe.

5. Average Period of Service of Employes.
6. Index numbers on rates of pay or earnings, employment or unemployment, hours of work, cost of living, and other important personnel facts for which they may be available. These data should be listed in detail for each month, quarter, and year covering the entire period of time for which they may have been worked up.
7. Accompanying computed or derived data there should be shown separately for each group of facts the following information:
   (1) Source of the group of data.
   (2) By whom the data were compiled.
   (3) The number and class of items included, particularly with respect to index numbers.
   (4) The method of compilation.
   (5) Sources of the detailed data from which the figures were developed.
   (6) The period of time covered.
   (7) For index series, the base period, if any.
   (8) The specific character of the data—*i.e.*, wages, hours of work, unemployment, etc.

The computed or derived facts listed above will be generally beneficial to all types of concerns, large and small, and a systematic attempt should be made to have them currently available, particularly if the primary external data noted above are not at hand. Even if the primary basic data are available these facts are serviceable by way of checks on comparable figures developed from the primary facts and the statistical organization should ever be on the alert to note, assemble, and properly record any such facts as may be of interest.

Historical External Personnel Facts.—Provision should be made for having at hand historical data in this field, covering such specific facts and such periods of time as may be desired. Historical data for the period covered by the basic records outlined above will be readily available from these records. No specific statement can be made as to what data should be developed historically beyond this and what need not be so developed. Special

problems may arise which will require historical development of almost any of the classes of facts listed above. Generally, for data not obtainable from the primary basic records, it is sufficient to make provision for historical data on the number of employes, absences, labor turnover and wages, handling such other historical data as may be required through special studies. The data listed under "Computed or Derived Facts" above usually are developed historically, particularly those which appear in the form of index numbers, and as such will frequently serve the need for external historical data of this type.

Future Estimates of External Personnel Facts.—Future estimates in the field of external personnel statistics will be required for different classes of data. The requirements for such facts cannot be fully anticipated and it is generally preferable to work them up by special study as occasion demands.

There are certain classes of facts, however, the future estimates of which may be more regularly required, such as wages or earnings, employment or unemployment, the number of gainfully employed, cost of living, immigration, and so forth. These data may be required for the business community or other territorial section and may be arranged according to sex, occupational classes, industrial classes, and in various other ways. Generally, computed or derived figures of this type will be adequate but the statistical department should be prepared to undertake special studies in the field whenever they are required.

Standard or Normal External Personnel Facts.—The development of standard or normal personnel facts in the external field is serviceable primarily as a basis for judging internal conditions. Consequently, the detailed discussion of this phase of external personnel facts has been included in the discussion of standard or normal internal facts on pages 23 and 24.

Additional Facts—General Information.—Many additional facts might be included in the list of external personnel data such as those listed below. It is believed, however, that the need for such facts is generally not sufficiently frequent or urgent to justify

comprehensive current files of them and that most of the data of this character can be adequately covered by special study as the need develops. In case the needs of the particular establishment require that any of these data be kept currently available, primary basic records may be developed for them. The list offered is not intended to be comprehensive but is designed primarily to suggest the general types of data that may have to be worked up from time to time.

1. *The Labor Situation*

Many additional facts may be recorded bearing on the general labor situation. Chief among these are the following which are typical of the group:

a. The number and characteristics of employes leaving the local labor market, distributed by their points of destination.

b. The number of employes entering the local labor market distributed by sources whence they come.

c. The number of gainfully employed distributed by sex, occupational classes, industrial classes, and territorial units.

d. Employment or unemployment facts distributed by sex, occupational classes, industrial classes, and territorial units. Also facts on unemployment insurance.

e. The number of strikes and lockouts showing the cause, the number of employes affected, the cost separately to the employes and employers, essential facts of settlement and so forth. These facts may be distributed by sex groups, occupational and industrial classes, selected territorial units and in such other manner as the needs of the particular situation may require.

f. Facts of labor legislation showing particularly the conditions imposed by law on business organizations, and facts reflecting the results of the operations of such laws.

g. Educational facts showing the relative number of grammar school graduates entering high school and high school graduates entering college, and the number of grammar, high school, private school and college graduates entering the industrial labor market.

2. *General Population Facts*

General population facts to be covered would include:

a. Internal movements of population between localities or different sections of the same locality.

b. The growth of population distributed in detail as may be desired, such as, for example, by sex, nationality, territorial units, and occupational groups.

c. Immigration facts showing the total number of immigrant aliens distributed by nationality, country of last permanent residence, destination in United States, sex, occupational class, age, and so forth. Also pertinent facts relative to emigration.

3. *Facts Relative to Organized Labor*

Many facts might be developed here, such as the number of trade unions, union hours and wages, union membership by broad occupational groups, showing both the actual number and the per cent to total employes in the group, and information regarding union agreements as they are made, showing the dates, number of employes affected, rates established, and other important items.

4. *Derived or Computed Facts*

Derived or computed facts would include facts relative to the health of industrial workers such as morbidity and mortality rates distributed by sex, occupational and industrial classes, territorial sections, nationalities, and age groups. Similar data may be required also for the general population.

(For references see page 79.)

CHAPTER V

INTERNAL FINANCIAL FACTS

Meaning of Financial Statistics.—In attempting to develop statistics on finances we must take care to distinguish purely financial facts from those relative to the financial phase of operations. Strictly, financial statistics are generally confined to the broad conception of the funds used to finance the business as distinguished from detailed items of operations such as facts relative to costs and prices.

From this fundamental point of view, internal financial statistics may be organized around three main heads—the acquisition of capital, the use of capital, and the return on capital. The facts offered below are developed on the basis of this three-fold classification. Here again, the details of each class will vary with the type and size of business and with the needs of the particular company.

I. THE ACQUISITION OF CAPITAL

A. *Present Securities, Notes and Mortgages of the Company*

1. For each item or class of items the following facts should be recorded currently:

a. Stocks—
- Number of shares authorized
- Number of shares issued
- Par value
- Date of issue
- Issue price
- Date listed on the Exchange
- Price history
- Yield history
- Dividend history
- Date of maturity or callable date

b. Bonds—
- Denominations
- Number of each denomination
- Date of issue
- Issue price
- Dates of interest payments

Interest rate
Date of maturity
Price history
Yield history

c. Notes—
Date drawn
Amount
Interest rate
Date due
Renewal dates

d. Mortgages—
Date drawn
Amount
Security—specific character, appraised value
Interest rate and dates
Date due
Renewal dates
Amount of principal paid if in gradual payments

2. *General Facts*

Sources of the capital obtained through the securities listed above as revealed by an analysis of the holders of them. This analysis should show such classes as the following:

Occupations
Geographical location
Amount held
Employes
Customers

II. THE USE OF CAPITAL

The classification of the capital which is in use in the business will be determined by the particular classes used in the accounting records. There will be a wide variation in the terms used. The classification below has been selected because of its simplicity. It probably would not fit any specific case precisely, but it will serve to indicate the type of data which should be developed in this field. These data should be prepared monthly and should show additions and deductions during the month as well as the total at the end of the month.

These facts will be developed regularly by the accounting department and furnished to the statistical organization. The

listing of the details from which they are developed does not fall within the scope of this discussion. More detailed facts of revenue and expenses and costs will be included in the list of operations facts presented in Chapter VII. All of the facts listed below will, of course, be expressed in dollars and cents and not in volume. From his study of accounting the student will be familiar with the terms used in the list.

A. Assets
- 1. *Capital Assets*
  - a. Intangible Capital—Organization
    Franchises
    Patent Rights
    Other
  - b. Physical Property—Real Estate
    Machinery and Equipment
    Construction Work in Progress
    Furniture and Fixtures
    Other
  - c. Investment Securities—
    Stocks of Controlled Companies
    Bonds of Controlled Companies
    Miscellaneous Stocks
    Miscellaneous Bonds
  - d. Other Capital Assets—
- 2. *Current Assets*
  - a. Inventories—
    (Raw materials, work in process, finished goods)
  - b. Working Assets
  - c. Marketable Securities—
    Stocks
    Bonds
  - d. Receivables—Notes
    Accounts
  - e. Cash— In banks
    In working funds
  - f. Special Deposits
  - g. Other Current Assets
- 3. *Deferred Assets*
  - a. Prepayments—Rent
    Taxes

Insurance
Advertising
Other

b. Other Deferred Assets

B. Liabilities

1. *Capital Liabilities*
   a. Stock— Amount by classes
   b. Funded Debt—Amount by classes
   c. Mortgages— Amount by classes
   d. Other Capital Liabilities
2. *Current Liabilities*
   a. Payables—Notes
      Accounts
      Dividends
   b. Other Current Liabilities
3. *Deferred Liabilities*
   a. Fixed Capital Reserves
      Accrued Depreciation—Plant
      General Equipment
      Amortization of Intangible Capital
      Other Fixed Capital Reserves
   b. Special Reserves
      Insurance and Casualty
      For Doubtful Accounts
   c. Miscellaneous
4. *Surplus*

C. Income and Expenditures

1. *Income*
   a. Primary or Operating Revenue
      Amount distributed by:
      Class of Products
      Plant or Company
      Territory
   b. Secondary or Non-Operating Revenue
      Interest
      Dividends
      Rents
      Other Secondary Revenue

2. *Expenditures*
   a. Purchasing Expense
      - Cost of Material
      - Salaries and Wages
      - Transportation and Storage
      - Other Purchasing Expense

   b. Producing Expense
      - Direct— Direct Labor
        Material
      - Indirect—Indirect Labor
        Factory Supplies
        Factory Expense
        (light, heat, power, etc.)
        Maintenance and Repairs
        Depreciation
        Other Indirect Expense

   c. Marketing Expense
      - Salaries and Wages
      - Commissions
      - Advertising
      - Credits and Collections
      - Traveling Expenses
      - Other Selling Expenses

   d. Overhead Expenses
      - General Expenses—Salaries and Wages
        Stationery and Printing
        Employes' Service Activities
        Finance and Accounts
        Other General Expenses
      - Taxes—Accrued
        Paid
      - Insurance
      - Departmental Administration

III. RETURN ON CAPITAL

These data will be prepared monthly or quarterly and will be distributed by plant or company, and by territorial units:

A. *Capital Used in the Main Operations of the Business*
   1. Gross Earnings
   2. Net Earnings

3. Interest and Other Deductions
4. Dividends
5. Balance to Surplus

B. *Capital Devoted to Non-Operating Activities*
1. Absolute Amount and Per Cent Return:
Securities of other companies
Loans
Sale of by-products
Rents
Services
Other
Total—Gross
Net

Historical Facts.—All important internal financial facts should be readily available historically by months, quarters, or years. Certain of these facts will be maintained currently in historical form and others will be worked up as occasion requires.

Historical facts in the field of internal financial statistics will be available for any period from the basic accounting records. For such data as may have been worked up regularly, historical facts may be obtained directly from summary records or financial statements and reports. For such material as may not have been so summarized historical data will be obtained from the basic detailed accounting records.

Future Estimates of Internal Financial Facts.—The facts included in the list below illustrate the type of internal financial data that need to be worked up for future periods.

I. BROAD FINANCIAL PLAN

The facts in this group will be worked up for both current or working estimates and long-time estimates. For the former group they will be more detailed and definite, but very careful attention should nevertheless be given to the long-time estimates.

1. Total amount of capital required distributed by:
Branches or companies
The time the capital is required
The uses to which it is to be put

2. Total amount of capital available from within the company distributed by:
   - Branches or companies
   - Time when capital will be available
   - Sources of the capital
3. Total amount of capital available through borrowings distributed by:
   - Branches or companies
   - Time when capital will be available
   - Sources of the capital
4. Total amount of capital required through new financing:
   - By branches or companies
   - By time when capital will be required
   - By the uses to which the capital is to be put
   - Proposed new financing to meet these requirements distributed by:
     - The form
     - The time
     - The amount

II. FUTURE ESTIMATES OF EXPENDITURES

Facts in this group will be developed in great detail for the period of the current or working estimate and in less detail for the period of the long-time estimate. The list of detailed facts given on pages 33 to 38 above should be pretty well covered in the current estimate. For the period of the long-time estimate only the broad groups, perhaps, will be required, such as purchasing, producing, marketing, and overhead expense.

III. FUTURE ESTIMATES OF REVENUES

For the period of the current or working estimates data will be developed showing the estimate of operating revenues distributed by classes of products, and territory, and the non-operating revenues distributed by detailed sources and in such other ways as the requirements of the particular company call for. For the period of the long-time estimate both classes of revenues will be shown only in total, distributed perhaps by branches or companies and by territories.

IV. FUTURE ESTIMATES OF EARNINGS

For the long-time period future estimates in this group will be confined for the most part to total net earnings, interest, divi-

dends, and balance to surplus. For the period of the current or working estimates the detail listed below should be provided.

1. Total net earnings
2. Total interest expense
3. Total dividends
4. Balance to surplus
5. Return on the investment—
   - Per cent total operating earnings to average plant in service.
   - Per cent total net earnings (operating and non-operating) to average capital obligation.

Standard or Normal Internal Financial Facts.—Systematic development and use of standard or normal facts in the field of internal financial statistics are all too rarely found among modern business concerns. Such standards are very effective instruments for quickly gauging the status and progress of the business. A complete and comprehensive list of standard facts will vary with different types and sizes of business establishments. The typical ones offered below are based on the "measures of efficiency" suggested on pages 50 and 51 of Bliss' "Financial and Operating Ratios in Management."[1] They are offered merely as illustrative: many others will be suggested by Bliss' textual discussion of these ratios.

Measures of Earnings—Standards on:
- The relation of surplus net profits to net worth.
- The relation of surplus net profits to volume of business.
- The earnings on common stockholders' investments.
- The relation of operating profits to total capital used.
- The relation of operating profits to volume of business.
- The relation of gross earnings to volume of business.

Measures of Costs and Expenses—Standards on:
- The relation of costs and expenses to volume of business.
- The relation of property expenses to investment and sales.
- The cost of borrowed capital.
- The cost of total capital employed.

Turnovers—Standards on:
- The turnover of capital employed.
- The turnover of inventories.

[1] Ronald Press Company.

The turnover of accounts receivable.
The turnover on fixed property investment.
Financial Relationships—Standards on:
The working capital ratio.
The manner in which capital is invested.
The sources from which capital is secured.
The proportion of earnings left in the business.

Additional Internal Financial Facts—General Information.—Additional internal financial facts may be developed by including facts which fall outside the scope of the outline given above or by listing more detailed facts within the scope of the outline. In the former group might be included such facts as the amount and character of banking business done, distributed by individual banks, and facts reflecting the company's credit. In the latter group may be added more detailed analysis of capital assets, receivables, inventories, liquid investments, payables, and return on the investment to reflect profit on individual classes or lines of goods. Such of these and other additional facts may be kept currently on the basic records as may be required in the particular case.

(For references see page 79.)

CHAPTER VI

EXTERNAL FINANCIAL FACTS

Need for Detail Limited.—The need for detail in the field of external financial facts will not be so great as in the field of internal financial facts. Frequently computed or derived figures will adequately serve the needs of the organization. Furthermore, the management of an industrial enterprise generally rely upon the advice of financial experts in important financial problems involving external facts, and therefore do not require as extensive detail as might otherwise be needed. These external data, though relating to corporate concerns, will nevertheless be of interest to the small non-corporate business.

I. THE MONEY MARKET
- A. *Money Rates*
  1. Commercial paper rates.
  2. Call money rates
  3. Time money rates
- B. *Security Yields*
  1. Yield on selected list of high grade bonds
  2. Yield on selected list of high grade preferred stocks
  3. Yield on selected list of high grade common stocks
- C. *Mortgage Rates*

II. THE SECURITY MARKET
- A. *Recent Corporate Financing (Stocks and Bonds)*
  1. Classification
  2. Amount
  3. Par or face value
  4. Issue price
  5. Rate of interest or dividend
  6. Yield
  7. Maturity
  8. Date of issue
  9. Source of capital if available

10. Methods of distribution (*i.e.*, direct sale to stockholders, employes, customers, or outside public, or sale to bankers or brokers)
11. Amount of issue absorbed

B. *Selected List of Securities in the General Market*

The detail listed under "a" and "b" on page 33 should be noted for each of the following:

1. General—Common stocks
   Preferred stocks
   Industrial bonds
2. U. S. Government securities
3. State and municipal bonds
4. Special list of securities comparable to those of the individual company
5. Securities of competitors

C. *General Security Market*

1. Stock Exchange transactions—Volume
   General trend of prices
2. Change in security prices as reflected by published averages, distributed particularly by industrial and railroad stocks, and municipal, state, U. S. Government, and industrial domestic bonds.

III. FINANCIAL FACTS OF OTHER COMPANIES

The facts included in the outline below may be listed separately for (1) industrial organizations in general, (2) the particular industrial group in which the individual company is classed, (3) certain selected industrial concerns which are of special interest to the individual company, such as potential or actual customers, concerns of which the individual company is customer for raw material, supplies, completed parts or finished goods, or organizations particularly essential to the operations of the individual company—such as railroads, steamships, telephone and telegraph companies, or (4) competitors of the individual company.

A. *The Acquisition of Capital*

The detail to be recorded in this group will depend on the interest of the particular company. For competitors and your own industrial group such statistics will be generally valuable. For each item or class of items the detail listed under "a" to "d" on pages 33 and 34 should be noted. In

addition more general facts relative to the number of stockholders distributed by classes of holders should be listed.

B. *The Use of Capital*

For each outside concern such of the balance sheet and income facts listed on pages 35 to 37 as may be required should be at hand. The particular classes of facts recorded will be governed to some extent by the classes that are used in the records from which the facts are taken. In any event, they will usually be broader and less in detail than corresponding internal facts.

IV. RETURN ON CAPITAL IN THE COMMUNITY

In addition to the facts of this general class which will be available from the data noted above for other concerns, those listed below will be serviceable.

A. *Regulatory Bodies*

Rate named as a reasonable return.

B. *Computed or Derived Figures* which purport to represent the general return on industrial investment in the community.

C. *The Net Income* reported on income tax returns related to the investment in manufacturing industries as reported in the Federal census of manufactures.

V. GENERAL FACTS

The external financial facts listed above have a more or less direct and immediate bearing on the problems of the business. In addition to these there are a great many important external financial facts which have a more remote and indirect bearing on the problems of the individual company. As illustrative of the type of such data the following groups of facts are listed:

A. *Banking Facts*

Banking facts will be listed separately for the Federal Reserve Banks and other banks. Out of many facts that might be listed in this group those outlined below are perhaps the most generally serviceable.

1. Federal Reserve System—

Total deposits
Total reserves
Ratio of Federal Reserve notes to gold reserve
Rediscounts—Amount
Rate
Notes in circulation

Due member banks
Bills on hand

2. Other Banks—Such of these facts as are pertinent should be shown separately for savings banks.
Resources
Liabilities
Deposits
Investments
Loans
Bank clearings
Debits to individual accounts
Cash in banks
Surplus reserves

B. *Money Facts*—
Amount of money in circulation
Velocity of circulation
Gross amount of money in the country
Price of money
Supply of money
Demand for money
Gold movements—Exports and imports
Domestic and foreign exchange rates
Capital movements
Fluctuation in the value of the dollar

C. *Public Finance*

Such of the facts listed below as may be of interest in the particular company will be recorded separately for municipalities, states, and the federal government.

1. Securities issued, distributed between those that are tax-free and those that are not. Details will be the same as those listed under "Recent Corporate Financing" on page 42.
2. Debts
3. Tariff duties
4. Internal revenue
5. Facts relative to income tax returns

Historical Facts.—Historical data should generally be kept currently for certain of the facts listed above, such as the following:

Money rates
Mortgage rates
Security yields
Stock market prices
Stock market averages
New corporate financing—amount, issue, price, yield, interest or dividend rate, etc.
Securities and return on investment of other concerns
Return on industrial capital generally in the community
Certain data listed above under "General Facts" such as bank clearings, debits to individual accounts, Federal Reserve ratio and reserves and rediscount rate, supply of money, and so forth.

The specific facts to be developed can be determined only on the basis of the definite requirements of the particular situation.

Future Estimates of External Financial Facts.—Generally, the management of an industrial concern will depend on the advice of financial experts relative to the future of financial conditions. There are certain facts, however, that should be estimated for future periods such as, for example, money rates, security yields, trend of stock market prices, demand for money, availability of capital. Generally, such services as are required from this type of data can be obtained satisfactorily from computed or derived figures that are currently available.

Standard External Financial Facts.—Certain general standards should be established in the field of external financial facts which will be serviceable in the development of standards of internal financial facts. The requirements for such external standards will vary with the type and size of the company. In general, however, the standard facts listed below will be desirable. These facts should be determined separately for general industrial enterprises and the specific group to which the individual company belongs.

1. Return on the investment
2. Relative amount or rate of depreciation reserve
3. Relative amount of capital in inventories

4. The rate of gross and net profits
5. The relative amount of working capital

These and other standards of external financial facts that may be required should be developed and revised as frequently as may be necessary.

Additional Facts—General Information.—Many additional facts in the field of external financial statistics might be listed such as facts on credit conditions, federal farm loans, acceptances, sources of industrial capital, financial facts of foreign industries, and so forth. Such additional facts may be required from time to time and in some companies it may be desirable to keep certain of them currently on file.

(For references see page 79.)

CHAPTER VII

INTERNAL FACTS OF OPERATIONS

List Comprehensive in Scope.—There will perhaps be a wider variation in the general character of the facts of operations among practical business concerns than of the facts of the personnel and finances. It is believed, however, that the outline offered below is sufficiently comprehensive to include the essential facts of the general type of large business organizations and sufficiently specific to be suggestive for the smaller establishment and the more specialized type of business.

It is not intended to fit specifically any particular concern. The facts required in manufacturing establishments, for example, would cover the entire list but those of a strictly mercantile establishment would probably be confined to the two broad classes, purchasing and marketing. Here, as in the case of the facts of personnel and finances, the list is intended to serve as a sort of broad basic starting point from which may be selected certain tangible facts that will be serviceable in the particular case. For purposes of this discussion the facts of operations will be organized under three main heads—purchasing, producing, and marketing.

Purchasing Facts.—The classification of facts under this heading is based on the different purchasing processes. The facts listed will apply to purchases of all kinds and classes of material such as supplies, or indirect material, raw or direct material, completed parts, or finished goods. The list of facts may be expanded for almost any type of concern and for certain special types additional groups of facts will have to be added. The facts listed should be recorded when pertinent respectively for each individual lot of goods represented by the separate requisitions, purchase orders, bills of lading, invoices, cancellations, receipt slips, issue orders, reserve orders, and returned goods orders; and also in connection with price quotations.

1. Date—All important dates such as the date goods are requisitioned, required, placed on order, shipped by vendor, received, etc.
2. Class of goods
3. Quantity of goods
4. Cost of goods—actual or estimated (in total or in form of unit price)
5. Use of goods
6. Vendor
7. Purchase or contract order number
8. Requisition number

In addition to the above, important facts peculiar to certain of the processes noted should be recorded, such as the person or department by whom goods are requisitioned, for whom they are reserved, and to whom they are issued from stores; supplemental purchase orders; class of transportation and charges in connection with goods shipped and received; reason and cost of cancelling requisition or purchase orders; in connection with goods received, the quantity accepted, and the quantity rejected showing the cause, disposition, and cost to the company; in the same connection also the source (purchase, return to store, cancellation of requisition or reserve) and disposition of goods received (stores, direct to factory, outside construction, direct to consumer); and the price submitted (list, discount, net, f. o. b. point) and dates effective in connection with price quotations.

Purchasing Costs.—Furthermore, facts on purchasing costs will generally be obtained from the cost department as required. Perhaps, the total purchasing cost distributed by classes of material and expressed in terms of cost per unit will be obtained currently. Additional detail showing the distribution of purchasing costs by detailed items such as freight inward, handling expense, supervision, and so forth, can be satisfactorily handled through special request to the cost department.

Miscellaneous.—In addition to the foregoing, many other facts may be included in this list such as those related to sources of material purchased, transportation charges, stock turnover,

and storage charges. Generally, these facts will preferably be assembled through special studies.

Producing Facts.—Production facts may be developed in almost unlimited detail. If we were to attempt to record separately for each individual product or group of products its complete history while in process, showing the time, material and labor required in each step, the detail would be tremendous. It is felt, however, that such extensive detail is not advisable on the basic records. A large part of this field should be covered by special statistical studies. The data listed here are designed to reflect the more general facts of the business.

A. *Production Orders*
   1. Date issued
   2. Date to be completed
   2a. Date begun
   3. Date actually completed
   4. Class of goods
   5. Quantity of goods
   6. Production order number
   7. Estimated cost
   8. Raw material— Received from stores
      Returned to stores
      In finished goods
      Waste
   9. Completed parts—Received from stores
      Returned to stores
      In finished goods
      Waste

B. *Shop Deliveries* (completed parts and finished goods)
   1. Date
   2. Class of goods
   3. Quantity of goods (separately for major parts and finished goods or classes thereof)
   4. Production order number
   5. Cost—Actual or estimated
   6. Completions
   7. Acceptances
   8. Rejections (number, cause and production costs)

C. *Shop Deliveries of By-Products*
1. Date
2. Class of goods
3. Quantity of goods
4. Cost—Actual or estimated
5. Value—Actual or estimated
6. Completions
7. Acceptances
8. Rejections (number, cause, production costs)

D. *Salvaged Material* (wastage, spoils, rejections)
1. Date
2. Class of goods
3. Quantity of goods
4. Cost—Actual or estimated
5. Disposition—Scrap, stores
6. Salvage value—Actual or estimated

E. *Unfilled Orders*
1. Date
2. Class of goods
3. Quantity of goods
4. Cost—Actual or estimated
5. Date required
6. Capital tied up
7. Capital available on completion

F. *Idle Machinery*
1. Department
2. Number of machine
3. Date idleness began
4. Date idleness ended
5. Cause of idleness
6. Normal output of the machine per unit of time (hour, day or week)

G. *Maintenance and Repairs*

The following facts should be listed separately for each repair or maintenance job:

1. Material consumed—Quantity, kind, cost
2. Labor—for each occupational class—
   Number of employes
   Amount of time
   Cost
3. Other cost items

H. *Production Costs*

The basic cost records will be kept in the cost department and the cost figures will be developed by that department. The statistical department should obtain currently from the cost department, data showing the total and unit cost for each kind of finished goods and completed parts produced. These data should be organized to show for each order completed the estimated or actual cost of:

1. Direct materials
2. Indirect materials
3. Direct labor—Regular time
   Overtime
4. Plant burden
5. Portion of general expenses

If the requirements of the particular situation call for it, the cost of material may be distributed by each kind of material component in the product, the direct labor cost by each class of labor, and the plant burden by its several detailed elements such as supervision, supplies, light, heat and power, depreciation, taxes, insurance, repairs and maintenance, and so forth. Certain items of plant burden may also be analyzed further to show the cost distributed by classes of labor and materials. All of these facts may be obtained separately for each step in the manufacturing process or each department of the factory.

The data listed above will generally be required separately for shop acceptances and rejections. Likewise, the salvage value of the rejections or the added cost required to perfect them should be available.

In addition to the actual cost data provided for above, corresponding standard cost figures should be at hand.

I. *General*

The statistical organization should have on file the following general information relative to production:

1. Complete list of all different classes of goods produced, distributed between finished goods and completed parts.
2. Complete list of materials, supplies and completed parts used in production.
3. Complete list of by-products produced.

J. *Plant Equipment*

The facts on plant equipment are listed under "Producing

Facts" for convenience, even though a part of such facts apply to other phases of operations. The details of such facts may be extended if need be but generally the data listed below will adequately serve the current need.

1. Machines—
   In this connection should be listed the total number of machines or groups of machines distributed by size, type and normal productive capacity. Provision should be made for recording important additions, deductions or replacements.
2. Building or Shops—
   Data in this group will generally be confined to the number of shops and the normal productive capacity of each. Here, also, important changes which have a bearing on production should be noted.

K. *Miscellaneous*

In addition to the foregoing there are many detailed facts which may be required from time to time such as those reflecting the production of different types of machines, the relative production of individual employes distributed by hours of the day and days of the week, wastage through fabrication distributed by individual processes, time and motion studies, changes in the character of the product, and so forth. Most of the detailed data of this type, however, will be developed by the administrative forces. Such of them as are required from the statistical organization should be developed as special studies.

Marketing Facts.—Facts on marketing should be very comprehensive and complete. The marketing of the goods is the crowning feature of the enterprise and really marks the potential if not the real limitation on the company's business. If sales stop, revenues fall off and profits cease. Some of these facts will be developed by the sales department and some by the statistical organization. As a general rule, the sales department should be relieved of the burden of developing all such facts except those required in the immediate operation of the department. The list of facts offered below is designed to indicate the data which should be available from the basic records of the statistical organization. Additions to, deductions from, and modifications of this list will

be made to suit the needs of the particular company. The list herein is organized to show detailed information in regard to the sales program and sales orders. The detailed facts required will vary with the number of salesmen, the number of different kinds of articles sold, and the extent of the territory. In addition to the facts listed below the statistical organization should have at hand a complete list of articles marketed distributed between those produced and those purchased, and high profit and low profit articles.

A. *Sales Orders*

For each individual sales order the following facts should be available from the basic statistical records:

1. Class, grade and size of goods
2. Quantity of goods—Total sold to customer
   Return allowance
   Total handled by customers from all sources
3. Date of sale
4. Place of sale (town, county and state)
5. Salesmen
6. Customer—Name
   Classification (*i.e.*, dealer, jobber, or private. The private group may be further classified, if desired, by economic, social, or racial classes or by such other classes as may be required)
   Old or new

6a. Lost Customers—Name
   Classification—same as for customers
   Reason

7. Selling price—Standard
   Quantity discount
   Trade discount
8. Cost price—Actual or estimated
9. Branch or company
10. Invoice number

10a. Production order number, if pertinent

11. Sales order number
12. Credit or cash transaction (with length of time extended on credit sales)

13. Delivery date promised
14. Actual delivery date
15. Cancellation—Date and cause
16. Goods returned—Date and cause
17. Complaint—Nature and cause

B. *Credits and Collections*

For each bill that is issued the following facts should be recorded:

1. Class, grade and size of goods
2. Quantity of goods
3. Date of sale
4. Amount of sale
5. Cash or credit, with length of time extended on credit sales
6. Date payment of bill is received
7. Collection costs
8. Uncollectibles

C. *Marketing Costs*

The detailed facts relative to marketing costs will be available from the cost department. Basic data records in this field will be kept by that department and cost figures developed as required. Data should be obtained from the cost department and kept on file in the statistical organization which will permit the classification of marketing costs including portion of general expenses by:

1. Territorial divisions
2. Salesmen
3. Classes of goods
4. Time periods—Days, weeks, months
5. Customers or classes of customers
6. Average per unit of product sold distributed by the sizes of sales orders
7. Branch or company
8. Old and new business

These relatively broad facts will generally be adequate for the needs of the statistical department. As occasion requires, however, these expenses can be obtained in greater detail, distributed, for example, by salaries, traveling expenses, postage, telegrams, express, and freight. Such details will rarely be required by the statistical department and they should be obtained as needed through special request to the cost department, no effort being made to obtain them currently.

D. *Advertising*

Facts in this field will be listed in three main groups—those related to the outlay or quantity of copy used, the results obtained and the cost. It is often very difficult to get definite final conclusions as to results of advertising. Generally, developments here should take the form of special studies. It is believed, however, that the facts listed below should be currently recorded so far as they are available.

1. Outlay or copy used
   a. Date
   b. Character of advertisement (such as catalogues distributed, magazine advertisements placed, newspaper advertisements placed, sample goods distributed, models distributed, posters, and bulletins)
   c. Quantity used or times repeated
   d. Location
   e. Class of goods advertised
   f. Class of customers appealed to
   g. Branch or company
2. Cost—Detailed figures will be worked up by the cost accounting department. Cost facts should be obtained and recorded to permit the following classification:
   a. By method used
   b. By territorial division
   c. By classes of commodities
   d. By time periods
   e. By companies or branch offices
   f. By class of customers appealed to
   g. By amount of sales and profits
3. Results—Accurate measurement of results is always difficult and usually impossible to obtain. Generally, the development of facts in this field will have to be done through special studies and should generally be accomplished in co-operation with the advertising department. So far as these are available the detailed facts of the results should be so recorded currently as to permit of the following classification:
   a. Total results obtained
   b. Results by methods used
   c. Results by territorial divisions

d. Results by classes of commodities
e. Results by time periods
f. Results by classes of customers
g. Results by companies or branch offices

The statistical organization may organize and develop the facts but the methodology in measuring advertising results should be determined in close co-operation with the advertising department.

E. *General*

Further facts in the field of marketing may be required such as the percentage we receive of each customer's total business, the number of calls made by individual salesmen, the results of the calls, and the cause of a failure to sell as determined by the salesmen; old customers from whom orders have not been received and the reasons therefor as reported by salesmen; adjustments on account of short weight, damaged or lost goods, and so forth; statistics of prospective customers showing the source of the prospect and the relative number who become customers; statistics regarding the business of customers such as the proportion of previous order on hand, to what extent he is stocked by competitors, his financial standing, the type of his customers, the general condition of his store and stock, and so forth. Much of this type of information, however, will better be handled by the sales department because of the significance of the individual fact rather than groups or summaries of facts. The work of the statistical organization will generally be confined to the handling of those facts which have to be organized into groups, classified and analyzed before their significance is apparent.

Historical Facts.—Historical data in the field of internal facts of operations are highly important. Some such data can be satisfactorily worked up through special studies as need for them develops. A large proportion of the data listed above, however, should be currently available in historical form by years, quarters, months, weeks or days as the case may be and in addition certain computed or derived figures should likewise be currently available in historical form. The following facts represent a minimum that should be kept currently on the basic records:

1. Purchasing—
   - Quantity purchased, by classes of goods
   - Cost—Price per unit
   - Total cost
2. Producing—
   - Shop deliveries, by classes of products
   - Cost—Per unit
   - Total
   - Raw materials consumed
   - Completed parts consumed
   - Salvage material
   - Unfilled orders
3. Marketing—
   - Sales—Quantity
   - Class of goods
   - Price
   - Sales deliveries—Quantity
   - Class of goods
   - Revenue—Collected
   - Uncollected
   - Uncollectible
   - Cost of collection per dollar of revenue
4. Advertising—Costs
   - Outlay
   - Results

These facts should be developed for such territorial divisions or other separate groupings as may be desirable in the particular company. In addition many other statistics will be worked up in historical form through special studies. These will include such facts as:

- Advance purchases of raw materials
- Completed parts or finished goods
- Price quotations
- Contracts executed
- Goods on special reserve
- Wastage in fabrication
- Sales cancellations

and such others as may be of interest in the particular case.

Future Estimates.—Future estimates of internal facts of operations will be considered separately for the period of the current or working estimate and the period of the long time estimate.

As has been stated before, the current or working estimate is confined to the immediate future and generally takes the form of a somewhat definite working plan. Basic records for this group will provide particularly for the following facts:

1. Purchasing Requirements—

   For each class of goods (*i.e.*, raw materials, supplies, completed parts, finished goods) the following facts should be at hand:

   - Quantity
   - Date required
   - Value—Actual or estimated
   - Purpose or use of goods
   - Unit price

2. Producing Requirements—

   For each class or line of goods the following should be available:

   - Quantity
   - Date required
   - Estimated cost
   - Estimated value
   - Production capacity of plant
   - Production costs

3. Marketing Requirements—

   It falls to the sales department to furnish the purchasing and production departments with the estimates of sales on the basis of which they can respectively formulate the purchasing and production programs. For each class of goods marketed, therefore, the following facts should be available from the basic records:

   - Quantity of goods required to fill the estimated sales demands
   - Shipments by months
   - Selling price, actual or estimated
   - Selling value, actual or estimated
   - Per cent cash sales
   - Per cent credit sales with an indication of the length of time credit is extended

Salesmen's individual quotas
Quantity that should be kept in stock
Revenue—Gross
Net
Selling costs

4. Construction Program—

This group will contain facts relative to the program for plant equipment. Such facts will be recorded here as may be considered important through their bearing on other facts of operations. It will generally not be necessary to keep extensive detail currently on record in this field. Facts reflecting any material increase in the probable output or productive capacity of the company will be kept currently at hand, such as material increase in the number of machines or shops, or the change in the type of machines to increase the capacity of the plant.

In addition to the foregoing such other current or working estimates will be developed as may be required in the particular case. They will be organized by months, quarters, years or otherwise as required.

Long-time estimates will contain less detail than the current or working estimates and will generally be developed for no more frequent periods than years. Long-time estimates should comprise future estimates of (1) purchasing facts, showing for each class of goods purchased, the quantity required, the unit price and the total cost, (2) producing facts, showing for each class of goods produced the requirements of shop deliveries, costs per unit and in total, and the plant capacity, and (3) marketing facts, showing for each class of goods marketed, the demand, the selling price, selling costs, and gross revenue.

Standard Facts.—Perhaps in no other phase of the company's activities are standards so serviceable as in the field of operations. Adequate operating standards properly developed will be of tremendous importance in measuring the results of the various operating activities of the company and gauging the efficiency of the operating organization.

The particular standards to be developed will vary with the type and size of the company. The list offered below is designed to illustrate the type of standard facts that will be required generally. It will be modified to meet the needs of the particular company.

Purchasing—

1. Amount of each kind of raw material that must be kept on hand to supply current needs.
2. Purchasing costs—Probably developed to show the costs per a certain selected unit.
3. Amount of each class of raw material required in special reserves.

Producing—

1. Quantity of raw material, supplies and completed parts consumed in the manufacture of a unit of output, distributed by the various classes of goods produced.
2. Shop deliveries per unit of factory equipment and labor force.
3. Per cent of total shop deliveries that are accepted.
4. Amount of wastage during fabrication, probably expressed in terms of the amount per unit of output.
5. Production costs per unit of output for each class of product distributed by material, labor, and plant burden.

Marketing—

1. Sales orders distributed by salesmen, territories, and classes of goods.
2. Percentage of sales orders that develops into sales deliveries for which revenue is actually received.
3. Amount of each class of goods to be kept on hand to meet current sales deliveries.
4. Profit per unit of product sold.
5. Percentage of total revenue, due on account of sales deliveries, that is collected.
6. Losses through customer's dissatisfaction with the product or service rendered.
7. Selling and advertising costs per unit, developed separately for the classes of product marketed and in such detail as may be required.
8. Marketing cost per dollar of revenue received.

The need for additional standard facts will develop in connection with the individual company. The foregoing are offered merely as suggestions and to emphasize the importance of having standard facts available.

Additional Facts—General Information.—Facts of this class have been noted on pages 49, 53, and 57, in connection with each of the three broad classes of facts outlined above—purchasing, producing, and marketing. The facts so listed are merely illustrative of the type of such data and are not presented as being comprehensive in scope.

In addition thereto there are many other facts that may be required from time to time such as data on management efficiency, the results of inventions or improvements of methods, facts on delivery methods, statistics showing the time elapsed between reporting and clearing maintenance trouble, progress of the individual order through the factory, and so forth. All such data can be developed through special studies.

(For references see page 79.)

CHAPTER VIII

EXTERNAL FACTS OF OPERATIONS

Wide Range of External Facts.—The amount of detail which might be worked up in the field of external facts of operations is almost unlimited. Here also greater use will be made of computed or derived figures and the basic data will require less development.

In selecting the facts listed below, the aim has been primarily to indicate the type and form of basic record data in this field and to suggest the actual facts which are perhaps most generally required in the majority of business organizations.

Like internal facts of operations, external facts will be classified primarily as purchasing, producing, and marketing or distributing facts, although certain special classes of data will be noted also, such as transportation facts, general price data, building facts, facts on business activity, and facts on general business conditions. Each of these three broad groups of data will be organized into two classes—*special* and *general*.

The *special* group will include those facts which have a more direct and immediate bearing on the problems of the particular business concern. These facts will be maintained currently on the basic records.

The *general* group will include those facts which are significant from the standpoint of business in general but have no special or particular significance for the individual company. Many of these data will be covered by special studies, the statistical department merely keeping informed as to where they may be obtained. In the former case, however, the regular type of basic records will not be generally required since the facts can usually be taken from the original sources—pamphlets, documents, bulletins, books—which will serve as the primary basic records of the statistical department.

External Purchasing Facts—Special Classes.—The facts listed below will apply to the different classes of goods particularly required in the operations of the individual business, whether in the form of raw material, supplies, completed parts or finished goods.

A. *Material Available—What the Market Offers*

Here again, the statistical organization will be interested only in the broad groups of facts and not in the individual items. The purchasing department will know where particular articles can be purchased at any given time. For example, the statistical records will show the amount produced, the quantity in transit, and stock figures, in general, say, for copper, textiles, or crude rubber, but these facts will not serve primarily as a basis for placing orders. The purchasing agent will be so familiar with the vendors and their stocks that he can determine at any particular time just where each kind of goods can be obtained most satisfactorily. For each class of goods the statistical organization will show the following facts relative to the goods available in the market:

1. Quantity
2. Location
3. Time
4. Class, grade and size

B. *Material Consumed*

Facts in this group will be confined to the industrial class to which the company belongs and to certain selected competitors. Within these limitations the statistical records will contain the following facts relative to the goods consumed:

1. Quantity
2. Time
3. Place (*i.e.*, territory or particular establishment)
4. General use
5. Class, grade and size

C. *Material Distributed*

Separately for each class of goods the statistical records should contain the following facts relative to the goods distributed by the main sources of supply:

1. Quantity
2. Source
3. Destination
4. Time
5. Class, grade and size

D. *Material Produced*

Facts in this group will be currently recorded in such form and to such extent as may be required in the particular case. A suggested list of data of this class is given below under "Production Facts."

It should be noted in connection with each of the four classes of facts listed above, that computed or derived figures will be available in many cases and should be utilized whenever they will satisfactorily serve the need. In such case, these facts may be transferred to basic records or the records may consist merely of references to sources where the material may be obtained.

General Purchasing Facts.—There are many facts which might be included in this group as having a general bearing on purchasing problems, such as data relative to transportation, business activity, general business conditions, exports and imports, and so forth. Most of these facts, however, will be included in the production and marketing groups given below.

External Production Facts.—

A. *Special Facts*

For the goods produced or marketed and raw materials, supplies, and completed parts used the following facts should be available relative to production:

1. Quantity
2. Time
3. Place (*i.e.*, territorial section generally and by certain selected competitors)
4. Class, size and grade

In view of the fact that the particular class of material or products will vary greatly with types of industries, no attempt is made to offer typical specific facts. For example, the telephone industry will be interested in such items as copper, lead, and cedar poles, the automobile industry in pig

iron, steel and rubber, the clothing industry in the textiles, and so forth. These will be covered more or less comprehensively in the section on general facts immediately following. Such of the facts noted in the general group as may be special for any company will be eliminated from the general group and treated specifically as outlined above. It is particularly important that the quantity of production be noted for those trades which make extensive use of the classes of products produced by the company. In addition to the facts listed above such data as are available on production costs should also be kept currently on the basic records.

B. *General Facts*

As was noted above, facts in this group serve to reflect the condition of business in general. They will be organized under three main heads—agriculture, mining and manufacturing. The list given here is designed to represent the type of facts that should be recorded and to indicate those which are most generally serviceable in business organizations of all types and sizes. It should be noted that certain of these facts will have a special interest for the particular company. These will be eliminated from the *general* list and included in the *special* group if the requirements of the particular case call for it.

Agricultural Facts—

Facts in this group will be confined to those which will most probably be of interest to all types of business concerns from the broad standpoint of general business conditions. Certain types of business concerns will have a special interest in certain agricultural facts. The meat-packing industry will be particularly interested in the production of live stock and the clothing industry in the production of textiles. These facts, however, will be included in the "special" group for such industries and will consequently be worked up in greater detail for them. This list of facts will be developed primarily from the standpoint of their bearing on general business conditions.

For various agricultural products, chiefly wheat, corn, oats, rye, cotton, flax, hemp, timber, and garden and

orchard products, the data listed below should be available from the statistical department. They will be kept currently on record or worked up by special study as may be required.

Acreage—(Separately for that planted and that matured)
- Time
- Class and grade of commodity
- Place
- Quantity

Production—Class and grade of commodity
- Quantity
- Time
- Place
- Yield per acre
- Total value in dollars

Likewise, facts indicating the production and value of dairy products, and live stock, particularly cattle, sheep, and hogs, are significant. These facts will be developed by such time periods and territorial units as may be required.

Mining Facts

Out of the rather extensive data that might be included, the facts listed below are offered as most generally significant in this field. Chief among mining products which may have a general interest are anthracite, bituminous coal, copper, zinc, lead, tin, iron ore and crude petroleum. For each of these products and such others as may be of interest in the particular company the following facts should be available relative to production:

- Quantity
- Class and grade
- Time
- Place
- Total value in dollars

Manufacturing Facts

Facts that may be listed in this group are many in number and type. The list, however, has been confined to those facts which have a more fundamental and far-reaching value and a more direct bearing on the general

business structure. Relative to the production of each of the following, there should be available

Quantity
Class and grade
Time
Place
Value in dollars:
Pig iron
Steel ingots
Copper smeltered
Automobiles
Automobile products
Domestic railroad cars—Freight
Passenger
Locomotives
Petroleum products
Textiles—Cotton goods
Woolen goods
Food—Wheat milled
Meat slaughtered
Sugar melted or refined
Canned goods produced
Electric power

This list might readily be extended to include a number of other items such as the production of boots and shoes, tobacco products, pulp and paper, cement, bricks, and paints. Data reflecting the volume of unfilled orders in certain classes of manufacturing establishments or for certain individual companies will also be important. This information should be shown separately for the industrial group to which the company belongs and those which are extensive users of the classes of goods the company produces. Likewise, general data reflecting the volume of manufactured goods may be of interest.

External Marketing or Distributing Facts.—

A. *Special Facts*

In this group will be included those facts which have a direct and important bearing on the marketing of the products of the business. The list given below will illus-

trate the general character of such data. For the goods marketed by the company the following facts should be currently available from the basic statistical records:

1. Consumption— Class, grade and size
   Quantity
   Time
   Place
   Classes of consumers—Economical, racial, occupational

2. Demand—
   The following facts should be developed separately so far as possible for actual and potential demand:
   Quantity
   Value
   Time
   Place
   Classification (*e.g.*, retail and wholesale, foreign and domestic, social, economic and occupational groups)

3. Competitors' Sales—Class, grade and size:
   Quantity
   Time
   Place
   Price

4. Custom Trades—
   Sales and revenues of trades which make extensive use of the classes of goods manufactured by the company and separately for important customers of the company.

5. Selected Statistics—
   In this group will be developed data particularly pertinent to the marketing of individual articles. The following are typical of this type of facts:
   a. Facts reflecting the number of residence houses electrically wired and the extension of electrical service, in connection with the marketing of electrical household appliances. In establishing the general potential market, these facts will be shown territorially for the

entire country. A more intensive analysis will be made where the interest is confined to a particular locality.

b. Data on the development of hard roads in connection with the sale of motor vehicles, accessories, gasolines, and oils.

c. Educational facts in connection with the sale of magazines or books.

d. Facts relative to the climatic conditions in connection with the sale of rubber footwear.

6. Additional Facts—

Many facts additional to those listed may be of special interest to the company, such as price data, internal revenue and income tax facts, exports and imports, population facts, individual incomes, and so forth. Such data are included in the *general* and *miscellaneous* groups immediately following. If the need require it, they will be included in the special list.

B. *General Facts*

General facts in the field of marketing and distribution may be grouped in a number of different ways, all of which might be equally justified. Data of this kind, typified by the list given below, will be kept on basic records or developed through special studies as the needs of the particular case require.

1. Wholesale Trade—Quantity
   Value
   Class, grade and size
   Time
   Price
   Stocks on hand by wholesalers
2. Retail Trade— Quantity
   Value
   Class, grade and size
   Time
   Place
   Counter or mail order
   Stocks on hand by retailers
3. Exports— Quantity
   Value

Class, grade and size
Time
Place (territorial section or particular port for both departure and destination)

4. Imports—
Quantity
Value
Class, grade and size
Time
Place (territorial section or particular port for both departure and destination)

5. Population Facts—

Detailed facts relative to the population in the territories in which the goods of the company are sold should be developed to meet the needs of the particular business. The specific character of such facts will depend on the type of article being marketed and the consequent character of the demand. For example, if the commodity is men's ready-made clothing, the population facts developed will exclude those relative to women and children. The list below is presented merely to indicate the general type of such data.

The total population of the territory being studied, distributed as follows, should be available from the statistical organization:

Male and female
Age groups—refined to suit the particular case, generally not more than three or four groups
Nationality or racial groups
Occupational groups
Income groups
Social classes

6. General Consumption Facts—

This group will contain facts relative to the consumption of various classes and grades of commodities which may have an indirect rather than a direct bearing on the sales problems of the company. The particular classes of commodities to be included in this group can be determined only on the basis of the needs of the particular concern.

Miscellaneous External Facts of Operations.—In addition to the foregoing external data there are certain groups of important facts which do not fall exclusively within any of the groups indicated but which are important to them all. The following are perhaps the chief classes of this type and they should be generally available from the statistical records:

A. *Price Data*

Facts in this group will take two general forms—basic facts and derived or computed figures in the form of averages, index numbers, relatives or percentages. The particular form to be used will depend on the needs of the particular establishment and the availability of the data. The records should show the data either as individual facts, average prices, index numbers, or percentage increase or decrease.

a. General wholesale and retail prices—

This group will contain general data designed to reflect the general price trend. Such data will usually be derived or computed in the form of index numbers, relatives, averages, or percentages.

b. Special wholesale and retail prices—

This group will contain wholesale and retail price data for the specific goods in which the individual concern is particularly interested either from the standpoint of material for manufacture or for marketing. These data likewise may be drived or computed. Specifically the prices of individual competitors may be listed here.

B. *Building Facts*

It is more difficult to secure accurate facts in this field than in some of the others. A great deal of estimating is usually done. Fairly good figures are generally available showing the facts listed below. These facts in the main reflect broadly and adequately for most purposes the activity of the building industry. For particular communities more detailed facts may be required which may be accumulated through special studies, or on current records.

a. Building permits granted— Time
Place
Value
Classification (*i.e.*, business,

residence, frame, brick, and so forth)

b. Building contracts awarded—Time
Place
Value
Classification (*i.e.*, business, residence, frame, brick, and so forth)

c. Building materials purchased by or used in the building industry. These facts should generally show for each class, grade and size of material the quantity, price, value, time and place. The following are suggestive of the building materials which should be included in this list:

Steel products
Cement
Brick
Tile
Timber
White lead
Copper

The facts may be listed for the various classes or grades of these several items if desired, such as Portland and other cement, fire-clay and face brick, oak and yellow pine lumber.

d. Contracts cancelled— Time
Place
Value
Classification (*i.e.*, business, residence, frame, brick, and so forth)

e. Miscellaneous—

Other facts might be added to the list such as those relative to building construction; house rents; real estate transactions showing the transfers, values, and general activity; taxation; restrictions on the use of land and buildings; and land and building values.

C. *Business Activity and General Business Conditions*

Certain facts listed heretofore tend to reflect business activity and general business conditions such as building permits and contracts, exports and imports, bank clearings, price data, and wholesale and retail trade. This group will

contain additional data which tend to throw light specifically on these points. The following data should be organized to show the volume or amount in each case distributed by such time periods and territorial sections as may be desired in the particular case.

a. Business failures—Showing liabilities
b. Car loadings
c. Sales of mail order houses and department stores separately—By classes of merchandise
d. Postal receipts
e. Advertising—By classes, such as newspapers, magazines, and catalogues

D. *Population Facts*

This group will contain general population facts regardless of any particular bearing on the problems of the business. Primarily these data will be confined to the total population distributed by such racial, age, sex, religious, educational, territorial, and other groupings as may be desired. Facts on immigration and emigration may also be included.

E. *Transportation and Communication Facts*

The facts in this group will be distributed by such time periods and territorial divisions as may be desired. They may also be organized to show certain sub-classes.

a. Railroads—
- Revenue passengers carried
- Revenue freight carried
- Revenue car loadings—
  - Grain and grain products
  - Coal and coke
  - Forest products
  - Manufactured products
  - Others
- Operating revenue
- Operating expenses
- Passenger rates
- Freight rates
- Idle cars

b. Steamship lines—
- Revenue passengers carried
- Revenue freight carried

Passenger rates
Freight rates

c. Express companies—Volume of business
Revenue from main operations by classes of goods carried

d. Telephone Systems—Number of stations
Number of exchange messages
Number of toll messages
Operating revenue—Total
Exchange
Toll

e. Telegraph Systems—Number of offices
Miles of wire
Number of messages
Operating revenue

f. Additional facts may be developed relative to transportation, by motor trucks and on canals, rivers, and inland lakes.

Historical Facts.—The facts listed below will be generally required in historical form. They will be maintained by such time periods and territorial groupings as may be desired.

1. Raw materials, Supplies, Completed Parts and Finished Goods, Generally Purchased by the Company
   a. Quantity available in the market.
   b. Quantity consumed separately by the entire country, by the industrial group to which the company belongs and the quantity exported.
   c. Quantity produced in the entire country and the quantity imported.
   d. Prices—Individual prices or derived figures particularly index numbers and averages.
2. Goods Produced by the Company—
   For each class of goods manufactured by the company, the following facts should be kept in historical form:
   a. Quantity produced separately by the entire country, and certain important competitors as well as the quantity imported.
   b. Quantity consumed by the entire country and also the quantity exported.

c. Prices—Individual prices or derived figures particularly index numbers and averages.

3. Competitors' Sales—

Quantity separately for entire country and a group of certain selected companies. Also individual prices of certain selected competitors.

4. Custom Trades—

This group should contain pertinent facts relative to the general conditions in those trades which make extensive use of the classes of goods manufactured by the company, such as the activity, volume of production, and sales and revenue. Such data should also be kept separately for large customers of the company.

5. Price Data—

In addition to the price data listed in 1, 2, and 3 above, facts relative to the general trend of prices should also be kept historically. Particularly index numbers of wholesale and retail prices.

6. General—

In addition to the foregoing, other facts may be maintained in historical form if the needs of the particular case require it, such as production of certain other manufacturing, agricultural, and mineral commodities; building facts; quantity of wholesale and retail trade by classes of goods; and general export and import facts.

Future Estimates.—Current or working estimates in the field of external operations will be confined to the immediate future and will cover such facts as the following:

1. Raw material, Supplies and Finished Goods, Generally Purchased by the Company for Use or Resale—
   a. Quantity produced
   b. Quantity available in the market
   c. Quantity exported
   b. The demand
   e. The price
2. Goods Produced by the Company—
   a. Total production separately for entire country and certain selected competitors
   b. The demand
   c. The selling price

d. Sales of certain selected competitors
e. Quantity exported
f. Quantity imported

3. Custom Trade—

General conditions in the trades which make extensive use of the classes of goods which the company produces. Specific data of this class should probably be developed separately for large customers of the company, both potential and actual.

Long-time estimates in this field will generally take the nature of future trends and will include such items as the following:

a. Trend of prices in general, wholesale and retail
b. Trends relative to goods produced by the company—
   Quantity produced
   Exports
   Imports
   Prices
   Demand
c. Trend of general business conditions
d. Trend of building activity and building costs
e. Trend of the general volume of exports and imports
f. Trend of production and marketing costs
g. Custom trades—Trend of general conditions in those trades which are large users of the classes of goods produced by the company.

Standard Facts.—The need for standards in the field of external facts of operations is quite as desirable as in other groups of facts. The selection of the specific standards, however, that will be required in any given company can be made only in the light of the detailed needs of the company. The standard facts listed below are offered as illustrative of the general type of such data. They will be developed separately for such time periods and territorial sections as may be advisable in the particular case.

1. Normal, annual usage in the country, of the classes of raw materials and supplies which are used by the particular company.

2. Of the various classes of raw materials and supplies used by the company, a standard reflecting the normal amount generally available in the market.
2a. Normal consumption of each class of the company's products.
3. Normal demand for each class of the company's products.
4. Normal production of the classes of goods manufactured by the company shown separately for the entire trade and for a selected list of competitors.
5. Standard amount of business done by the entire trade to which the company belongs.
6. Of the total amount of business done, the standard portion which the company gets.
7. Normal sales of selected competitors.
8. Custom trades—Normal activity for the trades which are large users of the company's products.

In addition to the foregoing many other standard facts may be required by the particular company such as normal building activity, normal exports and imports, normal business conditions, normal prices, normal production of agricultural, mining and general manufacturing products, and so forth.

Methods of developing these standard facts will vary with the type of standard. In some cases, derived or computed figures will serve; in others, developments will have to be made from a great body of detail.

Additional Facts.—There are many facts of general interest to a business concern which may be added to the list given above if required to serve the needs of the particular company. Typical of such facts are statistics indicating the wealth of the country as reflected in real estate, live stock, farm implements and machinery, railroads and their equipment, street railways, and so forth; fire losses, showing required detail; natural resources adequate in scope and detail to serve the needs of the company; results, from standpoint of business, of major inventions or improvements in business methods; additional taxation data and facts relative to foreign trade with particular reference to the balance of trade; pertinent facts relative to the produce exchange transactions; general information relative to the productive capacity of selected classes of manufacturing establishments in

the country; facts relative to government ownership and operation of industry; results of the operations of government commissions for the regulation of business enterprises; facts for foreign countries corresponding to those listed above; and facts relative to substitutes for the classes of products the company manufactures.

Generally, such facts can be satisfactorily handled through special studies. In case, however, that there is sufficient demand for any of them in the company, basic records can be readily developed. The final selection of facts to be carried currently on basic records will be made only when the needs of the particular company are clearly at hand.

References—Chapters II to VIII (Facts Required)

Bliss, James H. Financial and Operating Ratios in Management, 1924. (Financial and Operations Standard Facts.)
Duncan, C. S. Commercial Research, 1919, Ch. V.
Franklin, B. A. Cost Reports for Executives, 1913, Ch. VII.
Frederick, J. G. Business Research and Statistics, 1920, Ch. III.
Hall, S. Roland. The Handbook of Sales Management, 1924, pp. 201-209.
Haskell, A. C. Graphic Charts in Business, 1924, Ch. XX.
Jerome, Harry. Statistical Method, 1924, pp. 319-325.
Secrist, Horace. Statistics in Business, 1920, Ch. II.
Smith, W. H. Graphic Statistics in Management, 1924, Part III.

CHAPTER IX

STATISTICAL SOURCES—GENERAL

In Chapters III to VIII the discussion was devoted to the development of a suggested outline of detailed facts, internal and external, which will be generally required in working up a program of statistics in a modern business concern. We shall now turn to a consideration of statistical sources or where the facts outlined above may be obtained. This chapter will discuss certain introductory and theoretical matters of a more general character and those immediately following will indicate specific statistical sources from which the facts contained in the outline referred to above may be obtained.

Classification of Statistical Sources.—On the broad basis of proximity to original observations statistical sources may be classified in three main groups as follows:

1. Original Sources—The basis of all statistical data is a count, a record of observations made either incidentally in connection with other actions or independently. Regardless of the form in which data may come to the statistical organization, they initially represent individual observations of objects, actions, conditions, relationships, or what not. The original statistical source is the field of these observations. All data are obtained originally from this type of source. They can be collected directly from it only by making observations; no informant is involved.
2. Primary Sources—The primary source is the agent or agency that makes the original observations. The results of his observations may be furnished either from records as in the case of facts relating to his income, for example, or from memory as in the case of his own age or the ages of the members of his family. In any case, the original observation or count has been made and the data are being obtained directly from the agent who made them or from his records or publications.

It is important to note that in the primary source the data may be in any form. The observer may have them in original detail or in computed or derived form such as totals, averages, percentages, or ratios. So long as they are in the possession of the original observer they fulfill the requirements of the primary source.

On first thought, it may seem a strained distinction to say that the records of original observations are a primary source if they are in the hands of the agent who made them and a secondary source if in the hands of some one else. The distinction is very important, however, from the practical point of view. We want the experience of the observer as well as his records. It is often a great benefit to be able to discuss the data with the original observer regardless of their form.

3. Secondary Sources—The secondary source is an agent or agency that has obtained data directly or indirectly from the agent who made the original observations. Here again the form of the data is not an important consideration. The essential characteristic is that the agent furnishing the data did not make the original count. In this case the data may be obtained directly from the agent or from records or publications developed by him.

In brief we may say that the original source is the field of observations, the primary source is the agent who makes the original observations, and the secondary source is one which furnishes data that have been obtained directly or indirectly from the primary source. It is apparent that a single agency may be a primary source with respect to certain data and a secondary source with respect to others. A business concern, for example, is a primary source for all internal facts of the company whether they exist on basic records, have been compiled into reports or statements, or have been published in bulletin or pamphlet form for distribution to the public. On the other hand, if the business concern should collect data from a number of other companies it would become a secondary source with respect to these data.

This classification not only adds to the convenience of discussion but also is of practical benefit. From the standpoint of methods of collection and control over the character of the data,

it is distinctly serviceable. From the practical point of view, the methods of collection are generally distinct in the three types of sources although exceptions will be found with respect to the primary and secondary.

The responsibility for the character of the data is likewise distinct for the separate types of sources. In collecting data from an original source the inquirer is entirely responsible for the character of the data; in collecting data from the primary source only the inquirer and the original observer are involved; in collecting data from the secondary source, the responsibility is distributed among the inquirer, the original observer, the immediate informant and perhaps others through whose hands the data may have passed from the original observer to the immediate informant.

The significance of this classification from the standpoint of the collection of data will be given detailed consideration in the discussion of methods of collection on page 132.

Evaluating Statistical Sources.—As was indicated above, the statistical organization has only limited control over the data collected from primary and secondary sources, particularly the latter. It is therefore highly important to examine carefully any statistical data to be used from these sources. This will apply to both internal and external sources, although the fact that internal sources are generally official records of the company gives them a prima facie reliability which makes such thorough inspection unnecessary. No data collected from a primary or secondary external source should be used without the most painstaking examination beforehand. This will be especially important in the case of secondary sources which usually publish the data for distribution.

This position is not based on a philosophy of cynicism. There is no inclination to assume that everybody is either dishonest or simple minded or that all data collected by someone else are apt to be incorrect. The data may be perfect from the point of view of the purpose for which they were developed and yet not serve our purpose. The purpose for which data are collected plays such a tremendous part in the collection process that any data collected

for one purpose must be carefully reviewed before they are used for another.

Specifically, in examining any statistical source we should generally keep in mind the suitability of the data, their adequacy, and their reliability, for our particular purpose.

Suitability of the Data.—In examining the data which appear in any statistical source, a first consideration is their suitability for the purpose in mind. From this point of view we are interested, not in the reliability of the data nor whether the collection process was well done, but rather in whether the data in broader classification and in points of detail are of the kinds or types required to serve the end in view.

Data representing wages of factory workers may include all kinds of factories, whereas we may want the data for only certain types; they may include both male and female workers or all occupational classes combined whereas we may require the wages of male workers only or those of selected occupational groups. Prices of a certain commodity—wheat, cotton, copper, steel,—may be an average for several different grades or different quantities or different territories whereas we may want the price for a specific grade in a definite amount and in a particular locality. Sales data may reflect the orders taken whereas we may require shipments made or goods paid for. These examples will illustrate the importance of examining the data of any source for suitability to a particular need.

In evaluating the statistical source from the standpoint of the suitability of the data we may be guided generally by four aspects of the collection process by which the data were assembled—the purpose, the statistical unit, the character of the sample if only a sample study were made, and the method of disposing of specific difficult situations.

The importance of these various factors in the collection process is emphasized in Chapters XII to XX, beginning on page 130. Since they play such an important part in collecting the data they must likewise be taken into account in evaluating the data for any purpose other than that which guided the collection. If the data were collected for purposes of basic records, the aim will

have been generally to record comprehensively the detailed facts as is done in the population census of the United States. In such case the data collected will be suitable for a great variety of purposes. On the other hand, if they were collected for a specific purpose, the collection process will have eliminated everything which does not contribute definitely to the end in view. In this case the service of the data is much more limited. Regardless, however, of the character of the study, it is highly essential that the suitability of the data for the purpose in view be thoroughly established before the data are used.

Adequacy of the Data.—The data collected may be suitable in character and still be inadequate for the purpose in mind. The scope of the inquiry which produced the data and the completeness of the count within the scope selected are very important considerations.

For example, the statistical source may contain the kind of data we want but only for New York City whereas we may want them for New York State; for one or two business concerns whereas we may want them for many; or for quarterly periods of the year whereas we may want them monthly. In this case the kind of data available from the source is satisfactory but the amount is not sufficient.

The reference here is not to the sample study as compared with the complete induction. The sample may be truly representative and satisfactory in every respect but the field of data which it represents may be too limited in scope. The subject of sampling is developed in Chapter XXXIV.

Reliability of the Data.—The statistical source must likewise be checked closely for reliability. This will hold equally true for primary and secondary sources. What in general is the reputation of the agent who made the original observations? Is there a probability that he may be biased either consciously or unconsciously? Were the methods used in collecting the data proper and were they scientifically applied? If the source is a secondary one questions similar to these should likewise be satisfied with regard to agents, other than the original observer, who have handled the data.

The accuracy of the data should also be examined. This will involve a consideration of the accuracy aimed at and the accuracy attained. One purpose will call for a finer degree of accuracy than another and hence data may be satisfactorily accurate for the purpose for which they were collected and yet be not at all satisfactory for some other purpose. Information as to this point can be usually obtained from the explanations and directions accompanying the data.

We should likewise carefully note the methods of checking used in tabulations, classifications, and computations, and ascertain so far as possible what accuracy has been attained in these processes.

If the data represent a sample, as is generally the case in facts required by the business statistical organization, a thorough examination should be made of the sample chosen. We should determine primarily whether the samples taken were satisfactorily representative and the number of units sufficient to justify use for our purpose.

Furthermore, the time element in the collection process is also important. Were the data collected at a time when conditions affecting the observations were normal? The detailed discussion of the importance of this aspect of the collection process is given on page 138.

Actual Sources.—The next few chapters will be devoted to a discussion of actual sources from which the statistical organization of a business concern will generally obtain data. In this discussion the same general plan of development is followed as was used in the suggested list of required facts in Chapters III to VIII. Sources are grouped broadly into external and internal and within these broad groups they are, so far as feasible, considered separately for personnel, financial, and operations facts.

The general types of sources indicated on pages 80 and 81 are not considered as a fundamental basis in the classification. In the discussion, however, of each main group of sources we shall indicate whether it is generally original, primary, or secondary.

(For references see page 129.)

CHAPTER X

SOURCES FOR INTERNAL FACTS

Aim of Chapter.—Lack of adequate data is a very serious handicap in practical statistical work in a modern business concern. This is the case with internal no less than with external business facts. The number of business concerns are comparatively few which have complete statistical information relative to the fundamental phases of the management of the company, although such data are systematically recorded as a part of the normal business procedure. The purpose of this chapter is to indicate the records in the company on which the essential facts of the business are entered as a regular routine in connection with the functioning of the organization and which will therefore serve as sources of the internal facts outlined in Chapters III, V, and VII above.

The discussion of sources of internal business facts will therefore be confined almost exclusively to the consideration of records which are maintained for immediate and current use in the operations of the business. If the company has a reasonably complete system of personnel, financial and operating records, the detailed facts listed above can generally be obtained therefrom.

Primary Company Records as Sources.—The internal data contained in the primary basic records of the statistical department as outlined in previous chapters will be obtained chiefly from the primary company records. These records generally carry the individual facts of the business in original detail and are frequently records of original entry. They are consequently the most satisfactory sources of statistical information within the company and should be utilized extensively by the statistical department.

Secondary Company Records as Sources.—In some cases, however, resort to primary company records is uneconomical and unnecessary. The character of the data required makes it possible

to utilize secondary records or statements which are regularly available. In many cases secondary records contain the official figures of the company which is evidence that they have been carefully worked up and thoroughly checked. With reference, for example, to internal financial facts relative to the use of capital the statistical organization will depend almost exclusively on reports and statements prepared by the accounting department and will as a usual practice have no contact with the primary records.

In every case free use should be made of statements or records regularly prepared by the several departments for administrative purposes whenever the data contained thereon will serve the purpose.

Special Company Records for Statistical Purposes.—In rare cases is it necessary to develop special records for the primary purpose of furnishing statistical information. Generally, the records normally required in the administration or operation of the business are adequate. If the data required cannot be obtained from routine records of the company, it is generally indicative of a dearth of proper records and the method of procedure will usually be to introduce records which will serve both a statistical and an operating need rather than records to be used exclusively for statistical purposes.

It will be frequently necessary, however, to make certain additions on primary company records in order to complete the information required in the statistical service. These changes are generally only minor and can be accomplished with little cost and without in any way vitiating the records for operating purposes. If, however, the detailed facts do not approach those required in the statistical service, usually records can be substituted which will serve the operating needs equally well and also furnish adequate information for statistical purposes.

Association of Data.—Particularly, will it be necessary to make special provision whereby proper association of the facts may be possible. It is generally true that certain facts take on an added significance when associated with certain other facts

such as rates of pay by periods of service, age groups and occupational classes, or absences by sex, age, wage and period of service groups. It is primarily for this reason that provision is made for such entries as sex, age, period of service, occupational class, and so forth, on Figures 1, 2, 3 below. It is highly important that the association of facts be kept clearly in mind in establishing sources of internal business facts.

Evaluating the Sources.—There will be a great variety of primary company records distributed throughout the organization. Generally speaking, they will be equally reliable since for the most part they are, on the one hand, the basis for action in the functioning of the organization and, on the other, a record of the result of such action. In view of the fact, however, that their chief use is an operating one, we cannot assume their adequacy for statistical purposes. They may serve well the operating end for which they are intended and yet for different reasons not fulfill statistical requirements. They may be statistically incomplete as in the case of an Employes' Individual Record Change form which is to be used as a source of labor turnover data and which does not provide for such entries as sex, period of service and occupational class. Or the statistical unit may be unsatisfactory as for example when production statistics are based on completions rather than acceptances or sales statistics on orders taken.

There are many pitfalls in the use of internal sources even though they are systematically developed as a regular business procedure and are the primary operating records of the company. A careful examination should be made of those selected as statistical sources, keeping in mind particularly the points discussed in the previous chapter. Such an inspection will generally reveal the necessity, as was stated above, of altering the records in minor ways to make them more serviceable statistically.

Any secondary sources which may be used should be thoroughly examined in similar manner. In addition to the points noted above the methods used in arriving at any computed or derived figures in secondary sources should be carefully inspected. The comparative value of primary and secondary statistical sources in general is discussed in Chapter IX.

Location, Form and Designation of Company Records.—It should be emphasized that the particular designation of primary company records, their form, their location in the company, and the specific detail provided for will vary widely with different companies. This fact is fully appreciated in these discussions and the program of statistics presented herein has not been constructed on the assumption of any definite system of records. There is no inclination to assume that the primary records of a company are of a certain definite form or that the records required are to be found in certain definite departments of the company. The essential features of the program set forth here do not depend on any specific organization of the primary records of the company.

Furthermore, there is no inclination to suggest that the primary records of the business be recast generally to suit the convenience of the statistical organization. The position maintained throughout these discussions is that the outline of facts submitted can be generally obtained from primary company records as they exist in the particular establishment and that the necessity for any adjustment is negligible.

The program herein assumes only that the company has reasonably complete and comprehensive records of the essential facts of the business in one form or another and in one department or another, or, that, in the absence of such records they may be inaugurated with definite benefit to the company from an administrative as well as statistical point of view.

Illustrative Records.—The illustrative records presented below may, therefore, differ in many respects from the corresponding records in many companies. However, we are not primarily concerned with the location, form, and designation of these records. Such matters are of secondary interest only. We are fundamentally interested in their presence in the company and the detailed information which may be obtained from them. In selecting typical records, therefore, little attention was given to the size and the mechanical arrangement of the forms. It is sufficient for our purpose that the classes of records included in

the list exist generally among business concerns in some form and location and are very fruitful sources of internal business facts.

The specific illustrations offered are somewhat extensive in scope. They frequently consolidate in the one record a great variety of information which in many companies may be broken up and handled on separate and simpler records. Since our aim is chiefly to indicate general sources of internal facts and not in any way to treat the question of proper records, the more comprehensive form has been selected as facilitating the presentation of the subject from our point of view.

From his management course, the student will generally be familiar with the types of records listed. If not, supplementary readings should be assigned.

Survey of Existing Company Records.—In inaugurating a program of statistics in a modern business organization it is quite essential that complete information be at hand relative to existing records and the detailed facts contained on them. Before selecting records, therefore, to serve as sources of internal facts, a complete and comprehensive survey should be made of the important records of the company, both primary and secondary, listing (1) the names of the individual records, (2) the ways in which they are used, and (3) the detailed facts contained on them. On the basis of such a survey, the most serviceable primary sources for internal facts available from existing records will be determined. Furthermore, a comparison of this information with the data contained in the tentative outline of required facts, developed after the fashion suggested in Chapters III, V, and VII above, will indicate the adequacy of existing company records. What is wanted may be determined from the tentative outline and what is available from the survey.

Practical Aspects.—As a matter of practice, however, a statistical service in a business organization generally develops through gradual and systematic growth from small beginnings. Only the unsophisticated will undertake to work up an elaborate system of statistical records without first having carried certain selected phases of the program to completion. Proper methods of

procedure will lead first to the complete organization of certain facts that are readily available so that they may be at hand as required, such as wage statistics, security yields, sales data, and so forth. From this nucleus of actual service, the program will be gradually enlarged to include other classes of facts. In such case, the survey of existing records and modifications of them will generally take place simultaneously with the expansion of the program.

Presentation of Illustrative Records.—The types of records which may serve as statistical sources will be grouped into three classes—personnel records, financial records, and operations records. Since we are not interested primarily in the content of the forms, in some instances specimens have been narrowed disproportionately in the illustrations.

Personnel Records.—

Primary Personnel Records.—Personnel records may be considered as falling into three broad classes—engagement records, service records, and separation records. In other words, those containing facts relative to engaging employes, those containing facts relative to the service of the employes in the company, and those containing facts relative to the separation of employes from the company. Within each of these groups certain principal records have been selected to typify the various statistical sources. It is fully appreciated, however, that these principal records may be supplemented in all sorts of ways in any particular company.

The records listed below are offered as illustrative of the principal classes of personnel records in a business organization.

Engagement Records:

1. Application for Employment

Service Records:

2. Employes' Individual Record
3. Employes' Individual Record Changes (Figure 1)
4. Pay Roll Form (Figure 2)
5. Report of Absent Employes (Figure 3)

EMPLOYES' INDIVIDUAL RECORD CHANGES

COMPANY _____ DEPARTMENT _____ DIVISION _____

LINE No.	NAME	Sex	Age	Period of Service	CHANGE SERVICE RECORD AS FOLLOWS								DATE EFFECTIVE
						Payroll Code (1)	OCCUPATION		Dep't.	RATE OF PAY			
							Title	Code		Proposed Change	Maximum for Group	Basis (2)	
a	b	c	d	e	f	g	h	i	j	k	l	m	n
					From								
					To								
					From								
					To								
					From								
					To								
					From								
					To								
					From								
					To								
					From								
					To								
					From								
					To								
					From								
					To								

(1) The payroll code will generally indicate the classification of the employe as regular, part-time, temporary, office or factory. In case of new employe, enter the word "new" in column "g" opposite "From." When an employe leaves the service, enter in column "g" opposite "To," the code indicating the cause for leaving as shown on reverse of this form.

(2) D—Daily
W—Weekly
M—Monthly
Y—Yearly

Note—This space is used for approvals and for such directions as may be required.

FIGURE 1.—Employes' Individual Record Changes

This record is executed when an employe enters or leaves the service of the company or when his status is changed with respect to any important facts listed. It is used primarily to render advice on the basis of which the permanent record is currently corrected. It will serve particularly as a source for the data listed on pages 18 and 19 under "Labor Turnover Facts." Several simpler forms are frequently used instead of this consolidated one, such as a Transfer Slip, Salary Change Slip, Change of Title, and so forth.

EMPLOYES' PAYROLL FORM

COMPANY _____ DEPARTMENT _____ DIVISION _____

PAYROLL CODE (3) _____ WEEK ENDING _____

NAME	Occup. Code	Date of Engagement	Sex	Base Rate of Pay per Week		Regular Deductions				Regular Time							Overtime							Total Hours	Salaries and Wages		Special Payments		Special Deductions		Total Amount Due	Net Amount Due	Received Payment
						Stocks	Savings	Dues	Total	S	M	T	W	T	F	S	S	M	T	W	T	F	S		Regular	Overtime	Amount	Cause (1)	Amount	Cause (2)	(n+o+p)	(t−j−r)	
a	b	c	d	e	f	g	h	i	j	k							l							m	n	o	p	q	r	s	t	u	v

(1) Such as commissions, carfare, efficiency bonus, etc.
(2) Such as poor attendance, time lost, overpayment, etc.
(3) The payroll code will indicate the classification of the employe with respect to office or factory, and regular, part-time and temporary.

FIGURE 2.—Employes' Payroll Form

This record is used as a basis for paying employes. It is located in the payroll department and serves primarily as a source for facts listed under "The Wages" on page 19.

REPORT OF ABSENT EMPLOYES

_____ COMPANY _____ DEPARTMENT _____ DIVISION

DAY OR WEEK ENDING _____

NAME	Payroll Code	Sex	Period of Service	Occup. Code	Rate of Pay		Days Absent (2)							Distribution of Payments			Amount of Pay Deduction	Previous Excuse (3)	Cause of Absence (4)
					Amount	Basis (1)	S	M	T	W	T	F	S	Regular Payroll	Benefit Organiz.	Other			
a	b	c	d	e	f	g	h							i	j	k	l	m	n

(1) D—Daily
W—Weekly
M—Monthly
Y—Yearly

(2) Classify by means of following codes: (Combine codes as required—*e.g.*, "RF" indicates full pay chargeable to regular payroll)
R—Regular payroll
B—Benefit organization
O—Other
F—Full pay
H—Half pay
Q—Quarter pay
N—No pay

(3) E—Employe was previously excused
U—Employe was not previously excused.

(4) In sickness and accident cases, state the specific nature of case. Use reverse of this form if more space is required.

NOTE: This space is used for approvals and such specific directions as may be required.

FIGURE 3.—Report of Absent Employes

This record is prepared daily or weekly by the several departments or divisions of the company and forwarded to the personnel or industrial relations department or such other place as may be prescribed in the particular company. It serves primarily as a source for absence facts.

6. Report of Tardy Employes
7. Employes' Benefit Authorization Form

Separation Records:

8. Separation Form

Additional Sources.—Many records, report forms, and statements will be available in the particular company which have not been noted above and which may well serve as sources of internal personnel facts. Primary records in the respective fields will serve as sources for the facts listed under employes' Medical Service, Savings Plan, Cooperative Store, Restaurant Service, and Stock Plan. General instructions of the company will serve as sources for such facts as hours of work, vacation practices, holidays, and so forth. These additional records, together with the ones given above, furnish illustrations of types of sources which cover for the most part the entire list of facts outlined in Chapter III.

Still other personnel records are available generally in business concerns which will serve as statistical sources for certain kinds of data, such as Interview Record, Physical Examination Card, Mental Examination Card, Transfer Notice, Time Record, Rating Records, Requisition for Employes, and so forth. Secondary sources will also be serviceable in many cases.

The records noted herein are intended merely to typify the basic records of more or less general use. There has been no attempt to make the list completely comprehensive. As was emphasized above, it is very important at the outset to make a survey of existing records in the company and the ways in which they are used. On the basis of such a survey, the most serviceable primary sources for internal personnel facts will be readily determined.

Financial Records.—

Financial Records Standardized.—Owing to the fact that the financial records of the company are maintained primarily for accounting purposes, they are somewhat standardized in form, content, and location in the company. The detailed financial facts are assembled, classified, and summarized in accordance with a

generally accepted method of procedure, to produce the two final records of the company—the balance sheet and the income statement. All of this work is done by the accounting organization and consequently the financial records are located chiefly in this department. In the development of the data from the raw material to the two final statements, a definite system of principal records has been devised by the accounting profession which is followed more or less by all sizes and types of business organizations. Supplementary to these main records there will be a great number of others devised primarily as a means of getting the detail properly entered on the principal records. These supplemental records will vary widely with different types and sizes of business organizations.

Importance of Secondary Financial Records.—Since many financial facts of the business are currently organized, classified, and summarized as a regular accounting procedure, the various secondary accounting records become a very important source of internal financial data. The designation of the data contained thereon as the official company figures stamps them with a reliability which recommends their use wherever possible, and in no case should data used by the statistical department be at variance with them unless the disparity is fully explained and justified.

They cover practically the entire field of financial facts, particularly those relative to the use of and return on capital, and the bulk of important financial facts are generally available from them in forms satisfactory for most purposes. Typical forms of these records are given below.

Acquisition of Capital.—The facts listed in this group will be available from certain records located in the financial, auditing, accounting or treasury department as the case may be in the particular company. The details of these records will vary both as to form and content in different companies. Generally, the following will be available, although the specific designation listed may not be used in any given company:

1. Stock Subscription Blank
2. Stock Transfer Book (Figure 4)

STOCK TRANSFER BOOK

STATE OF NEW YORK

NAME OF COMPANY.

Date	Serial No. of Canceled Certificate	No. of Shares	By Whom Surrendered	To Whom Issued	Serial No. of New Certificate	No. of Shares	No. and Face Value of Stamps					
							2c	4c	10c	20c	50c	$1.00

(From Spurgeon Bell's "Accounting Principles.")

FIGURE 4.—Stock Transfer Book

This record is maintained in the financial department as the official record of stock transfers. It is a means of enabling the company to keep the list of stockholders currently correct.

3. Stock Register
4. Bond Register
5. Dividend Book (Figure 5)
6. Mortgage Register
7. Interest Detail Register

A comparison of the facts listed on pages 33 and 34 under the "Acquisition of Capital" with the details provided for on the typical forms will at once reveal the serviceableness of these records as sources of statistical information.

Use of and Return on Capital.—The facts listed on pages 34 to 38 relative to the use of and return on capital will generally be available from the two accounting records—the balance sheet and the income statement. While these records are somewhat standardized minor changes can be made to meet statistical

Hillsport Manufacturing Company

Semi-annual Dividend No. 24_____ 19____ 4%

Name of Stockholder	Address	Number of Shares	Amount of Dividend	We Hereby Acknowledge Receipt of the Accompanying Dividend

Figure 5.—Dividend Book—Personal Payment

This record is generally maintained in the financial department. It is a complete and comprehensive record of dividend payments made by the company.

requirements. The specific detailed items listed and the general order of arrangement will vary with different concerns.

Should greater detail be required, certain other accounting records will serve as sources. In addition to the General Ledger and General Journal, which contain some additional detail, there are certain subsidiary or auxiliary records which generally contain a great amount of detail and on which provision may be made for additional data that may be required for statistical purposes. Chief among records of this kind are the various registers or sub-journals such as the following:

1. Purchase Journal
2. Sales Journal (Figure 6)
3. Cash Journal
4. Notes and Bills Receivable Journal
5. Notes and Bills Payable Journal
6. Pay Roll Journal
7. Voucher Register (Figure 7)
8. General Cash Book
9. Bank Check Register
10. Collection Register
11. Sales Return and Allowance

In addition to the foregoing, similar types of records covering other classes of facts may be maintained. The accounting department will also generally prepare certain routine reports for administrative purposes. These are frequently fruitful sources of financial facts.

Operating Records.—Like the detailed facts listed in Chapter VII, the primary operating records will be grouped into the three broad classes—purchasing, producing, and marketing. The specific form of such records, their content, and the detailed ways in which they are used will vary widely with different types and sizes of companies. The forms listed below are typical of such records and will illustrate the kinds of data available from them. Detailed discussion of their use will be found in certain of the references listed on page 129.

Purchasing Records.—The records presented in this group will be applicable to the purchase of raw materials, supplies, com-

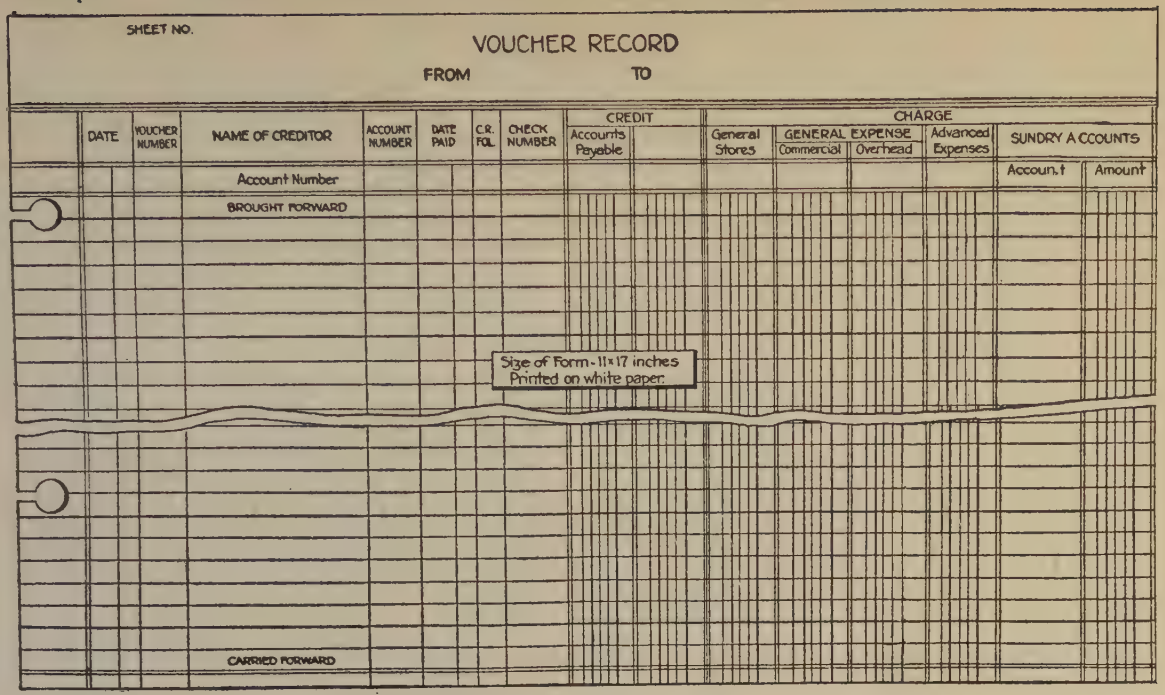
SALES JOURNAL

Date	Name of Customer	Address	Led. Folio	Invoice No.	Terms	Amount	Distribution *				
							A	B	C	D	E

* The letters appearing in the distribution columns represent classes of merchandise.

FIGURE 6.—Sales Journal

This record is maintained in the general accounting department as supplementary to the General Journal when the requirements of the particular company call for it. (Taken from P. J. Esquerre's "Accounting Theory and Practice.")



(From C. E. Woods' "Unified Accounting Methods for Industrials")

FIGURE 7.—Voucher Record

This record is generally maintained by the accounting department of all types and sizes although it will vary greatly as to form.

(COPY 1)

THE BUYING COMPANY

This number must appear on all invoices and packages
№ 35353

NEW YORK, N.Y. 192

ORIGINATOR OF REQUEST FOR MATERIAL

TO

REQUISITION NO.

PLEASE FURNISH US WITH THE FOLLOWING MATERIAL IN ACCORDANCE WITH YOUR QUOTATION OF

STORES SYMBOL AND LOCATION	QUANTITY	UNIT AND DESCRIPTION OF MATERIAL WANTED	LIST PRICE AND TRADE DISCOUNTS
		Return Acceptance Card Promptly, giving Complete Shipping Information.	

TERMS F.O.B.

SHIP VIA TO ARRIVE NEW YORK, N.Y.,ON

1. All packages must be marked with the order number given above.
2. Invoices and shipping receipts must bear this order number. Mail them on day of shipment.
3. Each shipment must be covered by a separate invoice. Articles on this order must not be billed with those on other orders.
4. The right is reserved to cancel this order if not filled within the contract time.
5. The conditions stated in this order shall not be modified by any verbal understanding or agreement.
6. Charges for boxing and carting will not be allowed unless previously arranged.
7. If price is not stated on this order, the material must not be billed at a higher price than last paid without notifying us in advance.
8. The acceptance of this order includes all terms, prices, delivery, specifications, and conditions stated herein.

THE BUYING COMPANY

PURCHASING AGENT

(From Madison Cartmell's "Stores and Materials Control")

FIGURE 8.—Purchase Order

This form is executed by the purchasing department when ordering goods. When supported by the vendor's acceptance it will serve as a primary source for facts listed on pages 48 and 49 relative to purchase orders.

PURCHASED MATERIALS RECEIVED REPORT COPY 1 ORDER NO.

RECEIVED FROM				DATE RECEIVED	DATE WRITTEN
VIA	CAR INITIALS AND NUMBER		NO. OF PACKAGES	GROSS WEIGHT	INVOICE NO.

MATERIAL SYMBOL AND LOCATION	QUANTITY			NET PURCHASE CHARGE	TRANSPORT. CHARGE	TOTAL COST	UNIT COST
	RECEIVED	REJECTED	ACCEPTED				
SYMBOL							
LOCATION	DESCRIPTION						
SYMBOL							
LOCATION	DESCRIPTION						
SYMBOL							
LOCATION	DESCRIPTION						
SYMBOL							
LOCATION	DESCRIPTION						

SIGNED	INVOICE CHECKED	TRANSPORTATION INVOICE CHECKED	CHARGES ENTERED AND COSTS FIGURED	ENTERED ON STORES CONTROL	
INSPECTION DEPT.					

(From Madison Cartmell's "Stores and Materials Control")

FIGURE 9.—Purchase Materials Received Report

This record is maintained by the receiving department and serves as a primary source for the facts listed on pages 48 and 49 relative to goods received. A receiving department copy of the purchase order may likewise serve the same purpose.

pleted parts, or finished goods. They will be classified according to the three principal phases of the purchasing organization—requisitioning goods, ordering goods, and receiving and storing goods.

REQUISITIONING RECORDS:

1. Purchase Order Requisition
2. Stores Requisition

ORDERING RECORDS:

3. Request for Quotations
4. Contract Order
5. Purchase Order (Figure 8)

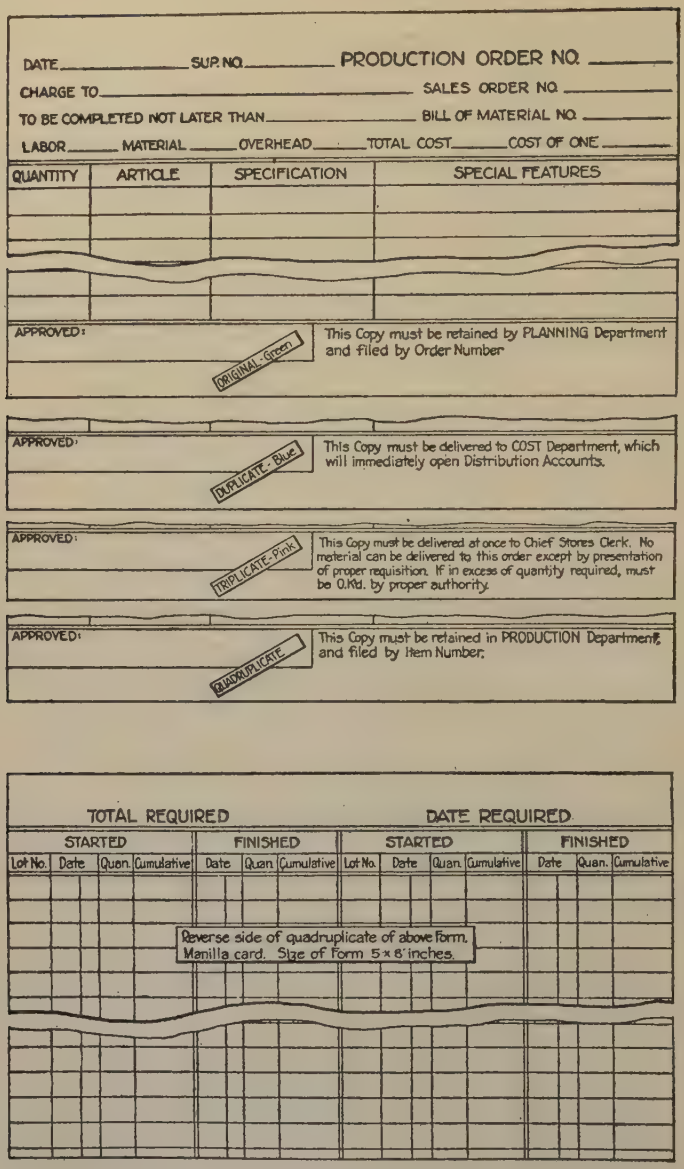
5a. Purchase Order Schedule

RECEIVING RECORDS:

6. Creditor's Invoice
7. Purchased Materials Received Report (Figure 9)

STORES RECORD:

8. Stores Record
9. Stores Return



(From C. E. Woods' "Unified Accounting Methods for Industrials")

FIGURE 10.—Production Order

This order is issued generally by the production department and serves as authority for production activity. Together with Figure 11 it will serve as a source of the facts listed on page 50 under "Production Orders," generally of those on page 50 under "Shop Deliveries," and sometimes for those on page 51 under "Salvaged Material."

Producing Records.—Records in this group will likewise be classified according to three broad phases of the production end of the business—production order records on the basis of which

ENGINEERING DIVISION
ROGERS MOTOR WORKS
BRIDGEPORT, CONN.
PRODUCTION SCHEDULE

NO._____ DATE_____

SYMBOL_____ TYPE_____ COMPONENT NO._____

TOTAL REQUIREMENTS_____
ON HAND_____ BASED ON MANUFACTURING ORDER NO._____
ON ORDER_____ RATE PER_____ SCHEDULE BEGINS_____
BALANCE TO SCHEDULE_____ NO. DAYS RUN_____ SCHEDULE ENDS_____

TIME PERIODS				SCHEDULES									
Week Ending		No Days		Schedule		Production Orders Issued				Made		Days	
MONTH	DATE	IN WEEK	CUM	FOR WEEK	CUM. TOTAL	NO	DATE		QUANTITY	FOR WEEK	CUM. TOT'L	AHEAD	BEHIND
SEPTEMBER	2	2	2										
	9	5	7										
	16	6	13										
	23	6	19										
	30	6	25										
OCTOBER	7	6	31										
	14	6	37										
	21	6	43										
	14	6	267										
	21	6	273										
	28	6	279										
AUGUST	4	6	285										
	11	6	291										
	18	6	297										
	25	6	303										
SEPTEMBER	1	6	309										

Size of Form - 7½ × 12 inches

(From C. E. Woods' "Unified Accounting Methods for Industrials")

FIGURE 11.—Production Schedule

This record is supplemental to the master schedule and provides for the parts and pieces of the models and units called for by the master schedule. It will serve as a source for data listed on page 50 under "Production Orders," those on page 50 under "Shop Deliveries," and those on page 59 under "Producing Requirements."

the production program is determined, the process records which have to do with the accomplishment of the program, and delivery records which have to do with the output of manufactured products.

PRODUCTION ORDER RECORDS:

1. Production Order Requisition
2. Production Order (Figure 10)
3. Manufacturing Order
4. Production Schedule (Figure 11)

PROCESS RECORDS:

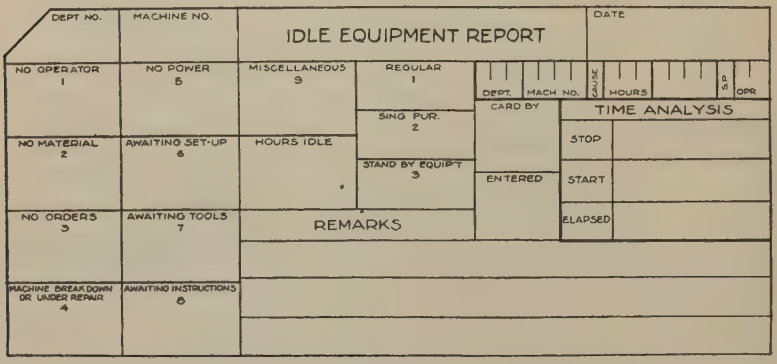
5. Idle Equipment Report (Figure 12)
6. Idle Labor Report

7. Repair Order
8. Betterment Order

Delivery Records—Total shop deliveries will be obtained from Production Order Records.

9. Process and Final Inspection Report
10. Rejected for Scrap
11. Rejected for Defective Repairs

Marketing Records.—The records listed below are typical of the records in this field which serve as sources of statistical data in this department of operations. Figures 13, 14, and 15—repre-



(From Madison Cartmell's "Stores and Materials Control")

Figure 12.—Idle Equipment Report

This form will be executed by the production department and will serve as a source for the facts listed on page 51 under "Idle Machinery."

senting respectively the Sales Order, Shipping Order, and Company Invoice—will illustrate the type of information which may be obtained from these records.

Many other operating records will exist in the particular company, such as the vendors' acceptance, bill of lading, cancellation order, and withdrawal order, in the *purchasing group;* the tracing tag or slip, production ticket, and time tickets, in the *producing group;* and the customer's receipt and advertising register, in the *marketing group.* These records are not standardized and in any given case it will be necessary to select those records as sources

of the statistical data required which will be most satisfactory and convenient. This is one of the most difficult steps in the development of a statistical program.

SALES ORDER

(Name of Company)

CUSTOMER — NAME_____
ADDRESS_____
CLASS_____
OLD OR NEW_____
SALES ORDER NUMBER_____
DATE OF ORDER_____
DELIVERY DATE REQUIRED_____
TERMS_____
NAME OF SALESMAN_____
TERRITORY_____

DATE OF SHIPMENT_____
INVOICE NUMBER_____
PRODUCTION ORDER NUMBER_____
WHERE SHIPPED FROM_____

Number Ordered	Description	Remarks	Price per Unit	Amount

Figure 13.—Sales Order

This form is executed generally by the salesman at the time an order is taken. It is used as a basis for shipment, for billing the customer, and may be used as a basis for issuing a production order. It will serve as a primary source for the facts listed on page 54 under "Sales Orders."

SHIP TO

CHARGE TO

REMARKS

STORES REQUISITION

COPY OF SHIPPING ORDER NO. 3694

F. O. B. | CUSTOMER'S ORDER NO.

TERMS | SHIPPING DATE

ROUTING

Location in Stores Room	Quantity and Unit	Material Symbol	Material Description	Unit Value	Extension

APPROVED BY

Order Department

MATERIAL RECEIVED BY

Shipping Division

Total Value

Apportioned		Issued		Deducted From Tag by	Deducted From Stores Record by	Extended	Checked	Distribution	Distribution
By	Date	By	Date						

FIGURE 14.—Shipping Order

This form is generally executed by the sales department or the shipping department. It will serve as a source for certain facts listed on page 54 under "Sales Orders." (Taken from Madison Cartmell's "Stores and Materials Control.")

THE SELLING COMPANY

New York, N. Y.

TO *The Buying Company* Date

New York, N. Y.

YOUR ORDER NO. OUR ORDER NO.

SHIPPED VIA

TERMS F. O. B.

Quantity	UNIT AND DESCRIPTION	List Price	Dis-counts	Net	Exten-sion

(From Madison Cartmell's "Stores and Materials Control.")

FIGURE 15.—Company Invoice

This form is generally prepared by the sales department or general accounting department. It will serve as a primary source for facts listed on page 55 under "Credits and Collections."

(For references see page 129.)

CHAPTER XI

SOURCES FOR EXTERNAL FACTS

Nature of External Sources.—The task of establishing sources for external facts is much more difficult than for internal. There is no definite system of records to which we may turn in the external field. With the exception of data obtained from other concerns, the more or less systematic list of sources in the statistical department will be built up to some extent in piece-meal fashion out of fragments gathered from a great variety of locations.

External sources generally have been developed by all sorts of agencies, in all sorts of ways and for all sorts of purposes, and are consequently greatly lacking in systematization, completeness, and definiteness of form and content. Frequently, the recording is done solely for the purpose of securing the facts rather than incidentally to the accomplishment of a certain end as is the case generally with internal data. In many cases, therefore, external records contain merely a group of facts developed through special effort at a given time and for a specific purpose and do not furnish a regular record, in complete and comprehensive form, of the essential facts of the activities or conditions to which they relate. Viewed from the standpoint of the total data available, the sources of external facts become one great conglomerate mass from which it would appear impossible to develop a systematic list of facts as a basis for a comprehensive statistical service in a business organization.

When properly approached, however, the problem is not so hopelessly difficult. While there is an enormous amount of data varying widely as to suitability, adequacy, reliability, and availability, nevertheless out of this wilderness of sources a great amount of valuable data can be selected, systematically recorded, and currently maintained on the basic statistical records. The proper method of procedure is first to determine in detail what

is wanted and then seek sources which will satisfactorily meet the requirements, rather than to undertake generally to organize the data which exist on external records. When this is done it will be found that adequate systematic information can be obtained from the external field to serve as a basis for a rather complete statistical program. While in many cases in the external field that reliability is lacking which comes from the official character of internal data, yet there are a number of external sources which likewise bear the official stamp of reliable agencies and can consequently be freely utilized in establishing basic statistical records.

Evaluating External Sources.—The nature of external sources makes it all the more important that the data contained in them be carefully examined for suitability, adequacy, and reliability in any specific case. Consequently, with each group of external facts listed there should be noted also certain information relative to their source on the basis of which we may evaluate the data for any specific use which may arise. Such information should include any facts which tend to qualify the data or limit their use or which in any way throw light on their serviceableness for the general purposes of the company. The detailed discussion of the importance and methods of evaluating statistical sources which is given in Chapter IX applies with special emphasis here.

Nature of the Discussion.—Owing to the nature of external statistical sources no attempt will be made generally to indicate the kinds of data available from each source or group of sources listed. Even the most successful efforts to classify statistical information available from the principal external sources have been serviceable to the practical statistical organization in only a very limited way. Such classifications are generally incomplete and only temporarily represent the actual conditions.

For the most part in this discussion we shall attempt merely to point out convenient methods whereby the statistical organization may keep currently informed or may readily learn, as to what data are available from the various sources considered. In connection with the discussion of each general source of data, there-

fore, specific methods will be suggested whereby the statistical organization may with minimum effort keep satisfactorily informed relative to the detailed facts available therefrom.

Other Business Concerns as Sources.—In considering other business concerns as sources of statistical information, it is worthy of note that the attitude of the larger industrial corporations toward the practice of exchanging information with one another has greatly changed within the last few years. The false impression of advantage to be gained through guarded secrecy as to the facts of the business is rapidly giving way in the face of enlightened self-interest and wholesome cooperation. Information which a few years ago was guarded most jealously is now handed freely to another company or even published for the benefit of the public. For instance, there is a general willingness to exchange detailed information on hours of work, labor turnover, and rates of pay whereas a few years ago such cooperation was almost unheard of. Because of this tendency other business concerns have become a most fruitful source for certain classes of external business facts. Let us hope that a continuation of this attitude may enlarge the scope of facts that will be readily exchanged until the entire group of personnel, financial and operations facts may be included.

It is generally feasible to obtain personnel facts and certain groups of financial facts from other concerns at the present time. It is suggested, therefore, that the most satisfactory source of the personnel facts noted on page 27 under "Primary External Data" will be a selected group of other companies. The financial facts listed on page 43 under "Financial Facts of Other Companies" may likewise be obtained from other concerns either directly or from annual reports and other published statements. Certain operations' facts may also be obtained from other companies in some cases such as certain of the special purchasing facts noted on pages 64 and 65, and certain production facts suggested on pages 65 to 68.

In undertaking to assemble data from other concerns care will be taken to select those particular companies which will best serve the purpose in mind. The group may include concerns which are

large users of the company's goods, concerns of which the individual company is a customer, competitors, or a general representative list of companies in the community.

Small Companies.—The foregoing section has been developed primarily with the larger type of industrial organization in mind. As was pointed out on page 15, statistical data may be of great service likewise to smaller business concerns and the cooperation among them in the matter of exchange of information will be quite as serviceable as among the larger corporations. It is perhaps true that along with the general lack of use of statistical data in the management of the smaller companies, there is also less inclination to cooperate by way of exchange of information. Such a practice, however, would be of great benefit. Through proper cooperation, moreover, the expense of such statistical developments would be relatively very small. The entire list of facts outlined in previous chapters will not be required. The facts chiefly required will probably be in the field of finances and operations. Within the scope of external facts pertinent in the particular case, efforts should be made to secure the information from a representative list of other companies. Closer cooperation along this line among smaller companies in a community will give them a certainty of management, will enable them to operate more economically, and will establish for them a firmer place in the business structure. The fundamental emphasis of this entire discussion will apply to the smaller as well as the larger type of company.

Federal Government as a Source.—The most extensive source of external statistical data is to be found in the several departments of the federal government. They are a veritable mine of information on a great variety of subjects in the fields of personnel, finances, and operations. Some of the data are collected currently or periodically and others at irregular intervals or even through a single special inquiry.

Lack of Information.—The data collected by the federal government are so many and diverse that it has been very difficult for the statistical organization of a business concern to keep currently informed relative to the detailed facts available therefrom.

To build up basic records in the statistical organization showing the detailed data is not generally feasible in this case. Added to the difficulty is the fact that there is no central body responsible for assembling or distributing statistical information but that several different governmental agencies are engaged independently in this work.

On this account there has been general confusion as to just what data existed on the records at Washington and the apparent difficulty of finding this out has greatly limited the statistical services of the federal government to practical business organizations. Those engaged in practical statistical work have often despaired of securing data from this source.

A Practical Method.—Efforts have been made by the government within recent years to overcome this difficulty and to make provisions whereby the public might more readily become informed as to what data were available from federal sources. The results of these efforts together with other developments have greatly relieved the situation. By putting forth reasonable effort at the present time, the practical statistical organization can keep satisfactorily informed relative to the data available from the federal government and the definite departments, bureaus, or other organizational units from which they may be obtained. This is accomplished by examining selected statistical publications that are issued and by keeping in touch with the general character of the statistical activities which are developed.

The practical statistical organization will determine by inspection what specific publications are sufficiently serviceable to justify their being kept currently at hand. These will vary with the character of the requirements of the particular company. Typical of such publications are the census volumes on population, agriculture and manufactures and the *Commerce Reports* and *Survey of Current Business* published by the Department of Commerce; the *Statistical Abstract* published by the Department of Commerce; the *Monthly Labor Review* and selected special bulletins such as those on Wholesale Prices, Retail Prices, and Cost of Living, Wages and Hours of Labor published by the Bureau of Labor Statistics; *Statistics of Income* published by the Treasury Depart-

ment; the *Federal Reserve Bulletin* published by the Federal Reserve Board; and so forth.

For current information beyond this, it will be generally sufficient merely to keep in touch with the broad classes of statistical activities. This end may be well accomplished through the federal government publications outlined below and the use of the mailing lists of the several departments and bureaus. Supplementing these general methods with special inquiries and inspection will generally keep the statistical organization well informed relative to data available from the federal government.

Report of the United States Bureau of Efficiency.—Perhaps the most helpful aid in determining the contents of federal sources generally is the "Report on the Statistical Work of the United States Government" issued recently by the United States Bureau of Efficiency under the direction of Herbert D. Brown. Though developed primarily for another purpose, it is of great value to the practical statistician who has occasion to use federal statistics. The purpose of the report as well as the scope of its content is indicated by the following taken from page 1 thereof:

> The object of this report is to describe the statistical work of each bureau and office of the Government reported upon, to point out duplications between offices where they exist, and to make recommendations looking to the improvement of the statistical product of the Government as a whole. No attempt is made to evaluate the quality, or to pass judgment on the technical aspects, of the statistics of any organization; such an attempt would lie outside the scope of the present report. The aim throughout has been to describe, not to appraise or criticise.

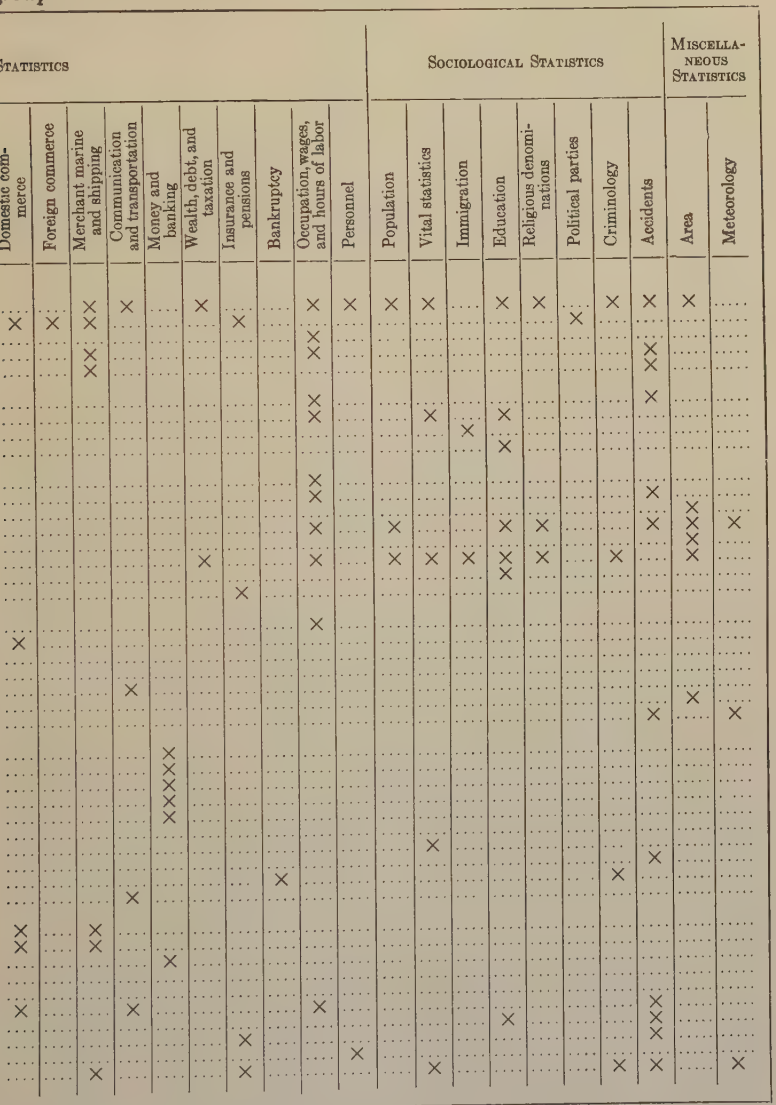
The appearance of this report has greatly increased the value of the various federal departments and bureaus as sources of statistical information. Its importance can scarcely be overemphasized. It has, so to speak, placed the great mass of federal statistics on a systematic basis, identified the specific location of each class of facts, and thus made these sources generally available with a minimum amount of effort. Whereas in the past we have been compelled to grope about rather hopelessly for statistical data in the federal government, we can now, with the aid of this report, pro-

Statement showing organization units that contribute statistical data to each can be

Department and Bureau	Economic						
	Agriculture	Mineral production	Forestry	Fisheries	Manufactures	Consumption and stocks	Commodity prices
Department of Commerce:							
Bureau of the Census	×	×	×	×	×	×	
Bureau of Foreign and Domestic Commerce						×	×
Bureau of Fisheries				×	×		
Bureau of Navigation							
Steamboat-Inspection Service							
Department of Labor:							
Bureau of Labor Statistics							×
Children's Bureau							
Bureau of Immigration							
Bureau of Naturalization							
Department of the Interior:							
Geological Survey		×			×	×	×
Bureau of Mines					×	×	
General Land Office							
Reclamation Service	×						×
National Park Service							
Office of Indian Affairs	×						
Bureau of Education							
Bureau of Pensions							
Department of Agriculture:							
Bureau of Crop Estimates	×					×	×
Bureau of Markets	×				×	×	×
Bureau of Animal Industry	×						
Bureau of Chemistry					×	×	
Bureau of Public Roads							
Forest Service			×		×		×
Weather Bureau							
Department of the Treasury:							
Office of the Secretary of the Treasury							
Division of Bookkeeping and Warrants							
Treasurer of the United States							
Bureau of the Mint		×				×	
Comptroller of the Currency							
Bureau of Internal Revenue					×	×	
Public Health Service							
United States Coast Guard							
Department of Justice							
Post Office Department							
Department of War:							
Corps of Engineers							
Inland and Coastwise Waterways Service							
Federal Reserve Board						×	×
Federal Trade Commission							×
United States Tariff Commission							×
Interstate Commerce Commission							
Federal Board for Vocational Education							
United States Employees' Compensation Commission							
Civil Service Commission							
The Panama Canal							

Figure

of the 27 subjects under which the statistical work of the Federal Government grouped



ceed with some degree of certainty to the particular departments or bureaus at Washington which are engaged in statistical activities of the general type in which we are interested. The accomplishment of Mr. Brown's recommendation relative to the establishment of a "Bureau of Federal Statistics" would greatly enhance the serviceableness of federal statistics. We should then have primarily only one federal source instead of the many as at present.

A copy of this report should be in the files of every practical statistical organization which may have occasion to use federal statistics. As Mr. Brown remarks in his notes on page 24 of the report, the statistical activities of the government will vary so that the description in the report as to detail will not long accurately represent the facts. However, it is probable that the main outline of the work suggested in the report will remain substantially correct so long as the statistical work of the federal government is done independently by different agencies. It is hoped that this report may be revised from time to time and that future revisions may be undertaken with the purpose of making the report serviceable to the public as a means of determining the facts available from the federal government.

Classes of Data Collected by the Federal Government.—In the report referred to above, Mr. Brown has grouped the entire statistical output of the federal government into twenty-seven classes. Figure 16 above is reproduced from page 2 of the report. It lists the twenty-seven groups and shows the various government organizational units which contribute to each group. Beginning on page 359, the report lists the classes of statistical data in much more detail, indicates the government organizational unit which develops them, and the frequency with which they are worked up.

These two sections of the report enable us readily to locate the classes of data which are available from federal sources. The detailed character of the data is treated in the main body of the report, which discusses the statistical activities by organizational units.

Department Lists of Publications.—In addition to the report referred to above. several departments of the federal govern-

ment put out pamphlets which list in detail, generally by subject groups and chronologically, all of the publications which they have issued. Typical of such pamphlets are the "Circular of Information Concerning Census Publications," "Publications of the Department of Labor," and "Publications of the United States Bureau of Labor Statistics." These pamphlets which are published at irregular intervals serve as valuable references in ascertaining in greater detail the statistical work of the several federal departments. Inquiries regarding them may be directed to the Government Printing Office at Washington, D. C.

Other Means of Keeping in Touch with Federal Statistics. Supplementing the use of the publications outlined above as a means of keeping in touch with federal statistical activities, the statistical organization should make use of the regular mailing lists of the several federal departments or bureaus and should resort freely to special inquiry. The publications referred to will be helpful in determining the particular government agency to which any special inquiry should be directed and in selecting the particular mailing lists which may be serviceable. These means will generally amply supplement the use of the pamphlets and enable the statistical organization to keep currently informed as to the statistical activities of the federal government. As was pointed out above some inspection and research will be necessary in determining the detailed facts within the broad groups determined from the method outlined.

Serviceableness of Federal Sources.—A great deal of criticism is heard of the statistical work performed by the federal government. Particularly is the business world inclined to belittle the statistical service which the federal government renders to the practical business organization. These criticisms fall generally into two broad classes—that the work is highly inefficient, resulting in a great deal of duplication, and that the data released are not sufficiently current to serve the purposes of the practical business concern.

While these criticisms may have some foundation in fact the situation is by no means as bad as some would have us think. In regard to duplication in the statistical work of the federal govern-

ment the "Report of the Bureau of Efficiency," referring to the table reproduced as Figure 16, makes the following statement:

> An examination of this table will show that frequently several bureaus and offices publish statistics in the same field. Accident statistics, for instance, are published by twelve bureaus, educational statistics by seven, and vital statistics by five. It must not be concluded, however, that the statistical work of the Government is characterized to any marked extent by duplication and overlapping. Where several bureaus publish statistics relating to the same general subject it will be found, as a rule, that they treat of different phases of the subject. The only cases of real duplication existing at the time of writing this report were the collection and publication of wholesale price statistics by the Bureau of Foreign and Domestic Commerce and by the Bureau of Labor Statistics, and the collection and publication of statistics relating to the production of oleomargarine and renovated butter by the Bureau of Markets and by the Bureau of Internal Revenue. Instances of apparent duplication are numerous, but except as here noted, they invariably prove to be cases where several organizations are engaged in the collection, compilation or publication of statistics relating to separate parts of the same general field rather than to the same identical subject.

This finding should tend to quiet those who so frequently hurl criticisms at the government for its supposed waste and inefficiency through duplication of statistical work. No justification is to be offered for even the kind of duplication that does exist, nor is the general method of handling the statistical work in the federal government to be endorsed. On the contrary it is fully appreciated that the collection by different agencies of statistical data in the same general field, even though the data differ in their specific character, may be highly inefficient and that the centralized statistical organization is undoubtedly preferable. Nevertheless there should be cooperation with the governmental agencies to the end that maximum service may be obtained from the data which exist.

As for the second criticism, it is a question whether that is not based on a false assumption of the proper function of federal statistical service. Generally speaking, the statistical service of the federal government is properly confined to the broad funda-

mental problems of business and does not apply to the current problems which arise from day to day and require quick decision. This latter type of service belongs to special facts accumulated by the individual business concern or by trade organizations. It is the broad economic facts rather than the detailed business facts which the government may well furnish. Government facts should be in the nature of research statistics; the individual business concern and trade organizations should furnish current material.

The great mass of data, therefore, which are available from the federal government will be serviceable primarily from the broad point of view. They will be used to determine broad questions of policy and not so much to accomplish the solution of current problems. For example, the business concern will use government statistics in establishing broad wage trends but it will resort to specific data collected from representative business establishments or other sources to reflect the wage standard of the community and to serve as a basis for the consideration of immediate wage adjustments. Regardless of what effort the government agencies may put forth to meet the demand for current service, we shall continue to depend primarily on other sources for current statistical data.

State Governments as Sources.—The various state governments of the country also compile certain kinds of statistical data. Although their statistical activities are much more limited than those of the federal government nevertheless in many cases they are valuable sources of statistical facts. The data on weekly earnings and unemployment published by the Industrial Commission of New York and wage statistics published by the State of Massachusetts will be illustrative.

State activities generally cover a variety of subjects and the facts are worked up more or less irregularly. Any attempt to classify the data available from them would be futile. In general no systematic statistical program is followed nor is there any material standardization of the work done by the several states. Some states do more statistical work than others and some develop one kind of statistics while others develop other kinds.

As a general thing, some personnel or labor facts, some

financial facts, and some operations facts will be found in each state government. The details will have to be determined by inspection in the particular case. The examination of publications issued by the state, the use of mailing lists of the state governments, and the resort to special inquiries will in most cases furnish satisfactory information relative to statistical data available from the various departments, bureaus, or other organizational units in the state government.

Municipal Governments as Sources.—The governments of the larger cities in the country are likewise fruitful sources of statistical facts. It is even more difficult for the data available from these sources to be classified, and for the statistical organization to keep currently informed relative to statistical activities in this field. Statistical work in municipal governments is generally characterized by irregularity and the facilities for informing the public are generally very limited. To determine the data available from these sources the statistical organization will have to depend for the most part on special inquiries, the examination of regular publications of the municipal governments, and the use of such mailing lists as are regularly maintained by the municipal governments.

Non-Governmental Agencies as Sources.—In addition to the foregoing there are a great many non-governmental agencies which compile statistical data of various kinds. These agencies are so numerous and differ so widely in character and location that it is much more difficult to determine the data available from them than from the governmental agencies.

In general, however, the practical statistical organization will direct its effort to keeping in touch with the statistical activities of these agencies rather than attempting definitely to classify the kinds of data which may be obtained separately from them. This method will apply to books and periodicals as sources, to educational associations, university bureaus of business research, private business concerns, and general commercial and industrial organizations such as chambers of commerce, manufacturers associations, merchants associations, civic associations, and so forth.

The specific methods of accomplishing this end will have to be

determined in the particular case. Ordinarily it may be accomplished by means of:

1. Regular publications of the agencies when such exist.
2. Reviews and notes on current publications which appear in certain periodicals such as the *American Economic Review, Journal of the American Statistical Association, Management and Administration, Journal of Accountancy,* and so forth. Also the "lists of additions" issued by libraries.
3. Special inquiries.
4. Special notification of statistical developments of the agency through regular mailing lists or otherwise.
5. Special arrangements with libraries or commercial service organizations whereby they undertake to keep us informed of all important statistical developments of a certain selected kind.

As was stated above, non-governmental agencies which may serve as sources of statistical information are numerous and vary widely in character. A comprehensive list of these agencies would require a volume. The broad classes given below are offered as suggestive of the different types of sources in this field.

Commercial Service Organizations.—Certain organizations maintained for commercial purposes undertake to furnish information in their respective fields for a stipulated fee. The service offered is generally of two types. First, the agency prepares and distributes regularly to its subscribers or customers certain information; second it undertakes to develop special studies on request from customers.

A great number of such services have grown up in recent years, covering all sorts of subjects. Some confine their statistical activities to the organization and interpretation of data which are readily available to anyone, while others undertake to develop original investigations. A rather comprehensive list of such organizations together with the general character of the service rendered can generally be obtained from well-equipped libraries. The list given below is offered as illustrative of this type of agency: as the author is more familiar with such agencies in New

York City, a large proportion of those listed are located in this locality.

1. Bloomfield and Bloomfield, Boston, Mass.
2. Industrial Information Service, Boston, Mass.
3. National Bureau of Economic Research, New York City
4. The American Chamber of Economics, New York City
5. Bureau of Applied Economics, Washington, D. C.
6. Public Affairs Information Service, New York City
7. Standard Statistics Company, New York City
8. The Business Bourse, New York City
9. The Commercial Digest, New York City
10. The Alexander Hamilton Institute, New York City
11. Prentice-Hall, Inc., New York City
12. Babson Statistical Corporation, New York City
13. R. G. Dun and Company, New York City
14. Brookmire's Economics Service, New York City
15. Bradstreet Company, New York City
16. Moody's Investors' Service, New York City
17. Poor's Manual, New York City
18. The American Institute of Finance, Boston, Mass.
19. Advertising Agencies

Industrial Organizations and Associations.—This term is used as a comprehensive one to include all those organizations which are more specifically designated as trade associations, commercial organizations, general research agencies, and so forth. Some of these organizations or associations, such as the trade associations, are created for specific purposes and their statistical developments are in the main confined to a certain field. Others, such as chambers of commerce, have a wider field and their statistical activities cover a greater variety of subjects.

There is a large number of such agencies in the country and their statistical activities vary widely both in character and extent. The following source books will be serviceable in keeping in touch with this type of statistical source:

"Commercial and Industrial Organizations of the United States," compiled by the United States Department of Commerce and distributed by the Government Printing Office, Washington, D. C. This pamphlet contains a great number, listing them by

broad groups such as (1) Interstate, National, and International, (2) State and Territorial, and (3) Local. The general character of the organization is indicated either by its name or by special notation.

"Classified List of Trade and Allied Associations and Publications in the City of New York," compiled by the Chamber of Commerce of the State of New York, located in New York City. This pamphlet lists the organizations and associations by subject groups. Similar publications exist for other cities.

"Personnel Research Agencies"—A pamphlet published as Bulletin No. 299 of the United States Bureau of Labor Statistics. It contains some 200 pages devoted to the discussion of Federal, State and Municipal, and Non-official agencies. This bulletin is of great service because of the informative character of the comments on the various agencies listed. It is distributed by the Government Printing Office, Washington, D. C.

A great number of such publications are available. Frequently, special lists can be obtained of organizations or associations which are devoted for the most part to a particular field. The statistical organization can generally accumulate a list of those whose activities are of particular interest and can arrange to keep informed relative to their statistical work. For data required in special studies we can readily identify in source publications, the particular organizations or associations which are promising sources and then by special inquiry or by the inspection of publications secure specific information.

The following are typical of industrial organizations and associations:

1. Chambers of Commerce—Federal, State and Municipal
2. Trade Associations
3. Merchants Associations
4. Bureaus of Municipal Research
5. National Civic Federation, New York City
6. National Consumers' League, New York City
7. National Industrial Conference Board, New York City
8. National Association of Corporation Training, New York City
9. Employment Managers Association

10. Retail Research Association
11. Organizations and Associations maintained by Organized Labor
12. American Railway Association
13. American Bankers Association
14. Bureau of Railway Economics
15. American Bureau of Metal Statistics
16. F. W. Dodge Corporation

University Bureaus.—In recent years many universities have established bureaus to do research work of various kinds. Typical of such are the business or industrial research organizations maintained by Harvard University, Carnegie Institute of Technology, New York University, Northwestern University, University of Pennsylvania, University of Wisconsin and others. Bulletin No. 299 of the United States Bureau of Labor Statistics noted above gives a list of personnel research agencies in colleges and universities. Well equipped libraries will generally be informed relative to the most important of such agencies. For the detailed character of the work done the statistical organization will generally have to resort to special inquiry and the inspection of publications. The university research bureaus frequently offer a service to the public for compensation and in such cases they keep their customers informed relative to their activities.

Educational Associations.—Various educational associations of the country serve to make available certain statistical data. Generally, such associations do not undertake statistical activities as associations but their several publications frequently contain serviceable data. The list given below is offered as typical. The official publication of each association is also listed.

1. American Statistical Association—*Journal of the American Statistical Association*
2. American Academy of Political and Social Sciences—*The Annals*
3. American Economic Association—*American Economic Review*
4. Academy of Political Science—*Political Science Quarterly*
5. American Institute of Accountants—*Journal of Accountancy*

These publications will be covered by such source books as the Readers' Guide and the Industrial Arts Index both of which will be found in libraries. Such of them, however, as may be of special interest in any particular case should be currently inspected.

Miscellaneous Agencies.—In addition to the foregoing more or less definite types, there are many miscellaneous agencies from which statistical data of various kinds may be obtained, such as the New York Stock Exchange and other similar organizations, banks and investment houses, particularly the Federal Reserve Banks, produce exchanges, boards of trade, business schools, employment agencies, boards of education, trade schools, religious organizations such as Young Men's Christian Association, Young Women's Christian Association, and Federation of Churches.

The character of the statistical activities of such agencies cannot be definitely outlined, although a general idea can be frequently obtained from the character of the organization. For example, stock exchanges will generally confine statistical activities to data relative to securities, while business schools, employment agencies, and religious organizations, will be serviceable in the field of personnel facts. Generally speaking, we can keep satisfactorily informed relative to such sources only through special inquiry.

General Publications.—A great deal of statistical information may be obtained from general publications such as periodicals, technical journals, newspapers, books and pamphlets. While the volume of literature in these various fields is very great, it is possible, with reasonable effort, for the statistical organization to keep fairly well informed with regard to sources of data.

Generally, the following methods should be used:

1. For periodical literature, technical journals, pamphlets and similar sources, the practical statistical organization will make free use of such aids as the Readers' Guide, Industrial Arts Index, 2400 Business Books and Guide to Business Literature, and Public Affairs Information Service. Such material as may seem interesting from the listings in these publications may be inspected as to detail.
2. For books, the statistical organization will find it generally

satisfactory to (1) inspect the lists of "additions" which are usually available from libraries, (2) examine the reviews and notes on current literature which appear in periodical publications and other places, (3) note the announcements of publishers of business books, (4) make use of card catalogues of libraries, and (6) examine reference books such as "Industrial Arts Index" and "2400 Business Books and Guide to Business Literature."

3. The data in newspapers are generally so definite and limited that little effort is required to keep in touch with them. They are usually confined to certain financial facts, price data, limited production and marketing information, and certain special types of facts such as rates of pay noted in advertisements for employes. The classes of data in newspapers do not vary greatly and the statistical organization can readily determine by inspection the data from this source that will be serviceable.

Original Enumeration.—Generally the sources indicated above will furnish adequate data for the program of statistics suggested by the outline of external facts in Chapters IV, VI, and VIII. In all of these sources the first step in the accumulation of statistical facts has been accomplished. The enumeration or counting of the individual units has been done and the data appear in a somewhat classified and summarized form.

In some cases, however, it may be necessary for the statistical organization to go back to first sources and make the original count of the individual units. This is particularly true in the field of marketing facts, where problems of marketing analysis arise. Other problems similar in character will arise in the field of marketing and to a less extent in connection with the other classes of operations facts and with personnel and financial facts.

No attempt to classify sources of this kind would contribute much of value to the practical statistical organization. The source required is generally indicated by the character of the data. The essential problem in this regard is the method of collecting the data, which is discussed in Chapter XIV. It is important, however, that the proper source be selected even in the original investigation, the question of time and place being particularly important.

References—Chapters IX to XI (Sources)

Bell, Spurgeon. Accounting Principles, 1921, (Financial Records).
Bloomfield, Daniel. Employment Management, 1920, Appendix 487-500, (Personnel Records).
Bowley, A. L. An Elementary Manual of Statistics, 1910, Ch. VIII.
Cartmell, Madison. Stores and Materials Control, 1922, (Operations Records).
Esquerré, Paul-Joseph. Applied Theory of Accounts, 1919, (Financial Records).
Jerome, Harry. Statistical Method, 1924, Ch. XVII.
King, W. I. Elements of Statistical Method, 1912, Sec. 20, p. 35.
N. Y. University. Source Book of Research Data, 1923.
Secrist, Horace. An Introduction to Statistical Methods, 1917, Ch. II, IV.
Tead and Metcalf. Personnel Administration, 1920, (Personnel Records).
U. S. Department of Commerce. Survey of Current Business—(Section on "Sources of Data" in the quarterly issues.)
Woods, Clinton E. Accounting Methods for Industrials, 1917, (Operations Records).

CHAPTER XII

COLLECTING DATA—GENERAL

Thus far in the general development of the subject we have considered what facts are required in a program of statistics and where they might be obtained. We shall now consider how they may be obtained or the methods of collecting statistical data.

Importance of the Collection Process.—It is frequently difficult to get the uninitiated fully to appreciate the importance of the collection step in statistical work. The collection of statistical data seems so commonplace and the task so simple that the amateur is inclined to look upon any comprehensive discussion of the subject as superfluous.

Contrary to this impression, however, the collection of data is one of the most important and in many cases one of the most difficult steps in statistical work. The detailed facts are the foundation of the finished statistical product. Collecting them is selecting the materials and laying the foundation from which the completed structure is to be built. During this process, the judgment is called into play constantly and decisions must be made which affect the entire statistical development. Frequently, the collection process is the most complicated and difficult part of a statistical study.

In connection with the development of index numbers of wholesale prices, the U. S. Bureau of Labor Statistics on page 25 of Bulletin No. 284, offers the following relative to the collection of the detailed data:

> The reliability of an index number obviously depends upon the judgment and the accuracy with which the original price quotations were collected. This field work is not only fundamental, it is also laborious, expensive, and perplexing beyond any other part of the whole investigation. Only those who have tried to gather from the original sources quotations for many commodities over a long series of years appreciate the difficulties besetting the task. The men who deal with data already

published are prone to regard all this preliminary work as a clerical compilation requiring much industry but little skill. To judge from the literature about index numbers, one would think that the difficult and important problems concern methods of weighting and averaging. But those who are practically concerned with the whole process of making an index number from start to finish rate this office work lightly in comparison with the field work of getting the original data.

We commonly speak of *the* wholesale price of articles like pig iron, cotton, or beef as if there were only one unambiguous price for any one thing on a given day, however this price may vary from one day to another. In fact there are many diffierent prices for every great staple on every day it is dealt in, and most of these differences are of the sort that tend to maintain themselves even when markets are highly organized and competition is keen. Of course varying grades command varying prices, and so as a rule do large lots and small lots; for the same grade in the same quantities, different prices are paid by the manufacturer, jobber, and local buyer; in different localities the prices paid by these various dealers are not the same; even in the same locality different dealers of the same class do not all pay the same price to everyone from whom they buy the same grade in the same quantity on the same day. To find what really was the price of cotton, for example, on February 1, 1920, would require an elaborate investigation, and would result in showing a multitude of different prices covering a considerable range.

Now the field worker collecting data for an index number must select from among all these different prices for each of his commodities the one or the few series of quotations that make the most representative sample of the whole. He must find the most reliable source of information, the most representative market, the most typical brands or grades, and the class of dealers who stand in the most influential position. He must have sufficient technical knowledge to be sure that his quotations are for uniform qualities, or to make the necessary adjustments if changes in quality have occurred in the markets and require recognition in the statistical office. He must be able to recognize anything suspicious in the data offered him and to get at the facts. He must know how commodities are made and must seek comparable information concerning the prices of raw materials and their manufactured products, concerning articles that are substituted for one another, used in connection with one another, or turned out as joint products of the same process. He must guard against the pitfalls of cash discounts, premiums, rebates, deferred payments, and allowances

of all sorts. And he must know whether his quotations for different articles are all on the same basis, or whether concealed factors must be allowed for in comparing the prices of different articles on a given date.

What has been said in this quotation relative to the collection of data for purposes of constructing an index number of wholesale prices will apply generally to the collection of statistical data of any kind. The complications of the task will vary, of course, with the character of the inquiry and the nature of the facts required. Generally, the task will be more difficult in the original and primary collection, but even in the secondary collection many complications are found.

Types of Collections.—Following generally along the lines of the classification of statistical sources given on page 80, collection processes may be grouped in three main types: the original collection, the primary collection, and the secondary collection. The original collection is the process of collecting data from the field of original observations, the primary collection is the process of collecting data from informants directly, and the secondary collection is the process of collecting data from publications. While theoretically this conforms in only a general way to the classification of sources, as a matter of practice it conforms pretty closely. The original collection is of course collection from original sources. The primary collection for the most part in practical statistical work will be confined to the collection of data from primary sources, although it may be used occasionally in the collection of data from secondary sources. Likewise, the secondary collection will be confined mostly to the collection of data from secondary sources but will be used occasionally in collecting data from primary sources. The closeness of this relationship from the practical point of view will be seen as the discussion of methods of collection proceeds.

PRELIMINARY CONSIDERATIONS

Thinking the Problem Through.—It is highly important that the person who undertakes the collection of statistical data should *think the problem through* at the very outset of his task.

Whether he is collecting data for basic records or for use in

a special study, in the internal or external field, or from original, primary, or secondary sources, he should thoroughly familiarize himself with every aspect of the problem confronting him so far as possible from information at hand. Literature on the subject should be inspected, particularly that containing accounts or descriptions of similar collections, and if convenient the problem should be discussed with persons who have had experience along the same line. On the basis of the information which may be assembled from all such sources the inquirer should give careful consideration to every phase of the proposed collection. He should try to visualize the entire undertaking and outline as definitely as possible the main steps in its development.

The engineer plans his structure so carefully that he can tell in minute detail the material that will be required at every step of the development, the processes that will be followed, and just how the complete structure will look. Likewise the statesman, the economist, the business man, all think their problems through before undertaking any sort of constructive work. This practice is equally important in collecting statistical data. Liberal time and energy spent in thoroughly considering every aspect of the problem at the outset of collection will pay ample return in both economy of time and money and in quality of results. It tends to establish the purpose and make definite the inquiry.

The Purpose.—The purpose for which the data are being collected is so important as to affect the entire collection. It serves as a guide throughout the collection process and is a serviceable test in the disposition of unforeseen difficult questions which arise during the actual work. The establishment of the specific purpose of the collection will enable us to avoid many difficulties that might otherwise arise and properly to dispose of those which do arise. If the purpose of the collection is not clear at the beginning and is not kept clearly in mind during its entire development, the disposition of particular difficult situations may be neither uniform in character nor correct for the purpose in mind. Failure properly to dispose of such difficulties when they arise will lead to serious troubles when the material collected is being classified or otherwise prepared for presentation.

If, for example, a study of wages is being made from the standpoint of the standard of living of the wage earner, we should require the data expressed in terms of earnings—*i.e.*, the total amount received by the wage earner; whereas if the purpose is to determine the wage level of the community we should want normal basic rates, eliminating deductions and special compensation, such as overtime, bonus and anniversary payments. If sales data are desired to reflect the relative activity of salesmen we should require to know the orders taken; if transportation costs, to know the shipments made; and if from the standpoint of revenue, to know the goods paid for. Likewise, in other classes of data the purpose is an important factor. It influences every phase of the collection.

From the broad point of view the purpose may be general or specific. In the former case the aim is the recording of all important detailed facts such as is done, for example, in the population census of the United States or the collection of internal business facts in the establishment of a program of statistics in a business organization. The purpose is to make the detail adequate for a complete and comprehensive system of basic records. In the latter, the aim is to collect only those facts which contribute to the specific end in view. In this case, the data must be carefully scrutinized both from the standpoint of general classification and of detailed character.

Defining the Inquiry.—One of the most fruitful benefits of thinking the problem through is that it makes the inquiry definite. This affects in an important way the time required in completing the collection and the quality of the results. If the inquirer has but a vague conception of what the collection involves, his course of procedure will be uncertain and his efforts consequently dissipated both in the scope of the survey and the character of the data collected. He cannot proceed effectively to the accomplishment of a piece of work unless he knows precisely what his task is.

The several steps in the development of the inquiry should be anticipated and the method of procedure made definite as far as possible. The scope of the inquiry, the statistical unit, the class of informants, the territory, the time of the collection, the methods

to be used, all of these and other important aspects of the inquiry should be determined. In a collection of wage data, for example, the inquiry should be sufficiently definite to determine the general industrial class or classes, the general occupational class, the territory, the time, the method, and so forth. All of these things may be varied in detail as the development proceeds but their general aspects should be clearly determined at the outset. The same will be true in the collection of any other kind of data. It should be emphasized that time spent in defining the inquiry at the beginning is time well spent, and invariably means time saved in the completion of the collection as a whole.

Practical Value of the Data.—The final test of the advisability of undertaking any given collection is the practical value of the data which it is proposed to collect. The business statistical organization is a service organization, designed to assist in the management of the business through the collection, analysis, interpretation and distribution of statistical data. It is consequently no more justified in collecting statistical data which do not contribute to this end than the production department of the business is justified in manufacturing products for which the company has no market.

It is very important, therefore, in considering the collection of statistical data that they be subjected to most severe practical test. Whether the collection is of a general character designed to furnish data for basic records or of a special type to furnish data for the solution of a specific problem, its ultimate service should be very carefully evaluated. This cannot be too strongly emphasized. No collection can be justified on any other basis.

Amount of Data Required.—In undertaking any collection, the amount of data required should be determined as definitely as possible at the outset. If the inquiry is for the purpose of collecting basic data to serve the statistical organization generally, the data will be comprehensive in scope. If the inquiry is for a specific purpose, the data required will be more limited. In any case, however, it is very essential that the amount required be determined before the actual work of collection is begun.

In this connection, the general principle of sampling developed

in Chapter XXXIV will be very convenient, which prescribes that essentially accurate conclusions can be drawn from a relatively small number of representative data. It is useless to collect a great mass of data when the same purpose may be served by relatively few.

Estimating the Cost.—It is a fixed law in business that no expenditure of money be made without adequate return. This principle applies to the collection of statistical data as well as to any other phase of the business. The statistical organization must be ever on the alert, therefore, to estimate the cost of any proposed collection and to match it against the anticipated value of the data to be collected.

SPECIFIC PLANS

The Statistical Unit.—In its essential aspects, the statistical approach rests fundamentally on a series of observations or a count. Whether we are collecting data from an original, primary, or secondary source; whether the data be internal or external; whether the data appear in their original detail or in computed form; in any case, they are the result of a count of individual units which has been made independently or incidental to some other action. Now the individual unit which forms the basis of any group of data we call the statistical unit. It is the thing which is counted. It may be a human being, a price, a house, a ton, a wage, a bushel, an acre, a vacation allowance, a rate of interest, a sale, a shipment, a receipt, an expenditure, or any of the many other items in the field of practical affairs.

It is evident therefore that this basic element is vitally important in any collection process. Before beginning actual collection the statistical unit or units must be clearly defined and these definitions must be carefully observed throughout the collection. This is apparently so self-evident that any emphasis given to it seems presumptuous. It is only a matter of common sense that we should know what is to be counted before undertaking to make a count or to assemble the results of a count of any given unit made by some one else. The problem, however, is not quite

so simple as it appears. A commonplace illustration will tend to clarify it.

What is a Book?—Suppose we were to undertake to count the books in a library. The problem seems simple. Everyone knows what a book is and the task merely requires that the books be taken shelf by shelf, counted, and entered on the record provided for the purpose. The actual development of a problem even apparently so simple as this, however, may develop complications which would require a more detailed and specific definition of a book than most of us have. Many questions arise which make the task a difficult one.

Is a paper-bound volume a book? Are pamphlets of 4, 5, or 6 pages to be counted as books? If Bryce's American Commonwealth appears in two volumes in one case and in one volume in another, is it to be counted as two books in the former and one in the latter? Is each volume of an encyclopedia to be counted as one book? If more than one copy of the same volume appears is each one to be counted as a book? Are those which are out of the library on loan to be counted? Are the multitude of government publications for which the library paid nothing to be counted as books?

When these questions and many similar ones that would undoubtedly arise in the actual count are presented, the definition of what is to be counted even in such a problem loses some of its simplicity and the importance of setting up a careful definition becomes apparent. In problems of more complicated character, the definition of the statistical unit becomes even more important and more difficult. Like difficulties will arise in defining an export; a price; a sale; an accident; a farm; a wage; or similar items.

Normal or Representative Conditions.—In planning any collection great care must be taken to see that the conditions represented by the data be normal or representative for the purpose in mind. This will involve particularly questions of the territory to be covered and the time the collection is made. It is not enough merely to collect a certain kind of data from a certain field of observations at a given time. An examination must be made of

the particular conditions within that field in which we are interested, to determine whether they are generally representative; and, if the field is generally representative, then further examination must be made to determine whether conditions are normal for purposes of the collection at the time it is to be made. Collection of wages of textile workers, for example, would hardly be undertaken from such territory as Illinois, Iowa, and Minnesota, or farmers' prices of corn and wheat from Maine, New Hampshire, and Vermont. Likewise, data on the normal production of an industry would not be properly collected from factories operating at only half-capacity. The question of time is particularly important also in collecting data relative to seasonal trades.

This question is important whether the collection be original, primary, or secondary. In the last two types, it will take the nature of examining the field of observations from which the data have been obtained. Information required in testing the conditions of collection may be obtained from literature on the subject, from interviews with those familiar in the field, or by making a preliminary inspection of the field of observations.

Accuracy.—The subject of accuracy may be considered in two aspects—accuracy *aimed at* and accuracy *attained*. Both of these aspects of accuracy are important in planning for a statistical collection. We must determine what accuracy we want in the particular data being collected and must provide adequate checks to give us reasonable assurance that this accuracy will be attained.

Providing methods designed to assure this accuracy is a problem to be handled separately, for the most part, in planning each individual collection. Distinct methods will frequently be required for each of the three fundamental types of collection. A discussion of the various detailed methods of checking accuracy in statistical work is given in Appendix E.

Selecting the Type of Collection.—Entirely different methods of procedure are required in the use of the original, primary, and secondary types and the expense in time and money also varies with them. Consequently, it is important that we know at the outset which is to be used.

Broadly speaking, this involves the determination of where

the data are to be obtained. In practice, however, it involves particularly an examination of secondary sources or publications of primary sources to determine whether the data required may be available in published form. Such an examination may reveal data which have been worked up for the same general purpose or those which have been collected for a different purpose but which will fulfill the requirements of the proposed collection.

If the data required in the proposed collection are available from publications, the secondary collection will be followed. Otherwise, the original or the primary collection will be required, according to the facts of the particular case. Since the comparative expense in time and money is generally much less in the secondary collection, it is very important that thorough search be made for the material in secondary sources before undertaking to obtain it from original or primary sources.

Survey of Similar Collections.—In the search for material already available special attention will first be given to similar collections—that is, collections which have undertaken to assemble a group of data of the kind it is proposed to collect. This is very important even if there be little or no hope of finding the required data. When a man contemplates building a house or purchasing an automobile he first inspects many houses already built and examines many varieties of automobiles. He is endeavoring to profit by the experience of others and build the best house or purchase the best automobile he can for the money invested. Likewise, in undertaking a statistical collection the same plan is to be followed; similar work done by others should be surveyed.

In addition to the possibility of obtaining the required data, the survey of similar collections will be particularly valuable by way of furnishing helpful suggestions in connection with the original or the primary collection. This is especially true in the case of general problems in the field of business statistics and particularly among business organizations in the same general industrial class. Problems are in the main similar in character and investigations made by one concern in a certain class are usually valuable to another concern of the same sort and oftentimes answer a specific inquiry that the second concern has

raised. Consequently, it is advisable to make inquiry among organizations which might logically be interested in the particular problem proposed to determine whether a collection which has already been made may be of some value in the proposed collection.

It is a common practice in business management, when a question is up for solution in one company, to find out how other companies are handling the matter. This tendency to take a broad-minded, cooperative attitude is characteristic of the business world to-day. Hence, it is especially advisable for one who is planning a statistical collection in the general business field to ascertain so far as possible what other companies have done along the same line.

Methods of Collection.—Likewise, before the actual collection is undertaken, the specific method of procedure should be determined. This method will vary with the general type of collection, separate methods being required generally, for the original, primary, and secondary types. Within each broad type of collection the particular method to be used will vary with the nature of the inquiry. The detailed methods generally serviceable with each type of collection are discussed fully in the next few chapters.

(For references see page 210.)

CHAPTER XIII

COLLECTING INTERNAL FACTS

This chapter will be devoted to a discussion of the methods of collecting the detailed internal data outlined in Chapters III, V, and VII. As was pointed out above, the collection of internal data consists primarily of transcribing them from the carefully kept company records. While this is a relatively simple process compared with the making of an original count or of collecting data from the more complicated external sources, nevertheless there are many difficulties in its accomplishment.

The Problem.—As already emphasized, the scientific consideration of the important problems of the business will necessitate detailed facts throughout the entire scope of the three fundamental phases of the organization—personnel, finances and operations. The great mass of detail involved, distributed as it is throughout the organization and located on so many different types of records which have been designed primarily to serve an operating end, makes the problem of assembling the facts a difficult one. The problem is further complicated by the fact which is emphasized throughout these discussions that the statistical service should be developed with minimum burden on the operating departments. In spite of the fact that detailed data are systematically recorded as a part of the normal business procedure, to transcribe them from the company records and put them in shape for efficient statistical service is no small undertaking.

In general, the methods of accomplishing this task may be grouped into two broad classes: (1) those in which the transcription of the facts is done by the operating departments either in response to special requests or through the preparation of routine reports; (2) those in which the transcription of the facts is done by the statistical organization either from the files of the operating departments or from copies of operating records maintained exclusively for statistical purposes.

Operating Departments: Special Requests.—In this case the statistical organization usually has no immediate control over the actual work of collecting the data, this work being done by the operating departments of the company. When the need develops, the statistical organization requests the data required from the proper departments or other organizational units. From the standpoint of the statistical organization this method consists of little more than issuing a request for certain material. The request may take the form of an interview, a telephone conversation, a memorandum, a statistical schedule, or a statement form. Further work of assembling the material is done by the operating departments, which actually transcribe the data from the company records. The statistical organization will generally, however, keep in touch with the preparation of the data which have been requested.

In general, the questions of method are confined largely to stating clearly just what data are wanted and the way they should be organized. Such a problem does not usually involve the delicate and difficult questions present in the methods to be used in collecting data from the more complicated external sources or in an original enumeration. Within the company, a reasonable interest and a fair spirit of cooperation can be assumed. Furthermore, the facts are taken from the official records of the company and will be generally compared by the executives in a number of different ways. On this account, the operating department will generally take great pains to see that the data submitted are accurate and truly represent the facts. No one in the company fancies having the executive return to him, for correction, data which he originally submitted as being accurate.

While this method may be satisfactory in some cases, there are, nevertheless, serious objections to its use. It places the real burden of collecting or transcribing the data on the operating departments, the irregularity of the requests being an especially disturbing feature. Furthermore, it requires too much time for the statistical organization to get hold of the required data to enable it to render efficient service.

Operating Departments: Routine Reports.—Internal statistical data may be assembled from the several departments of

the company by developing a system of routine statements which the departments prepare and forward periodically to the statistical organization. These statements will be of two general types. One type contains primary data and is designed to serve the statistical organization with basic information from which a regular statistical program may be developed; the other type contains certain derived or computed data—summarizations, associations, computations—and may be distributed by the operating departments to the statistical organization or directly to the management. These routine reports are supplemented by special reports as required.

The details of this method, from the standpoint of the statistical organization, consist primarily of the preparation of the statement or report forms which serve as the medium for transferring the data from the primary records of the company to the statistical organization on the one hand and the management on the other. The work of actually transcribing the material is done by the operating departments.

This method is generally preferable to the one just discussed. It nevertheless places definite limitations on the statistical service. The regular service of the statistical organization is confined to the character of data which appear on the routine statements received from the several operating departments. Any inquiry which involves additional data or the same detailed data organized in a different way generally necessitates a special request.

In the matter of final routine reports to executives the situation is somewhat different. Those which contain data of a single department can be handled very well in this manner, the statistical organization developing the form of report and the operating department currently preparing and distributing it. Generally, however, routine reports involve the combination of the facts of two or more departments, and in this case the data must be assembled from the several sources of the company and recast into a general report. Hence the statistical service to be obtained from final routine reports prepared in this manner is likewise limited.

Furthermore, the routine statements will be adequate for an extensive statistical service only in case they are very detailed in character and comprehensive in scope. They are in reality the basic records of the statistical organization and to be satisfactory

must contain the detail outlined in previous chapters. The preparation of such elaborate reports, however, will not only heavily burden the operating departments but will also be uneconomical, since a great amount of the detail will not need to be transcribed regularly from the company records if these records are readily available, as is the case when a system of primary basic records is maintained. Because of this fact, practicable routine reports are generally limited both in scope and detail. For the most part, they are inadequate for a satisfactory statistical service and particularly so in the preparation of statistical data which involve the combination of the facts of the several departments of the company.

However, good statistical service may be rendered with this method of assembling internal facts. It is worthy of more consideration when the operating departments of the company are scattered over a wide geographical territory or in the case of a holding company collecting facts from its several subsidiary or associated companies. Whenever it is used, it will be much more satisfactory if statistical operating units are established in the several departments to handle statistical work exclusively. The statistical organization can generally arrange to have these operating units keep certain data currently at hand and thus avoid the delay which might otherwise occur.

Statistical Organization: Use of Operating Copies.—In some instances the statistical organization undertakes to assemble the internal statistical data by transcribing them from copies of the company records which are used by the operating departments. The transcription is made at the time and place where it will least inconvenience the operating departments. The records may be used by the statistical organization after they have been finally filed by the operating departments or at some point of their flow through the organization. The problems involved in this method are generally confined, therefore, to determining the point at which to tap the flow of records through the operating departments and of selecting the files of operating records to use. When these have been determined the data will be transcribed to patented card records, report forms, Hollerith cards, or other statistical records.

The most effective use of this method requires that a small

statistical force be located in each of the operating departments, to transcribe the data when the records are not in use by the operating departments. These employes should be under the direct supervision of the statistical organization at least from a technical if not an organizational point of view. They can then be developed into a force which is intimately familiar, on the one hand, with the particular operating records with which they have to deal, and on the other hand, with the statistical point of view, and thus render more efficient service from the standpoint alike of output and accuracy. The work of transcribing the facts for both basic statistical records and special studies, being under the general supervision of the statistical organization, is done uniformly in all departments and from the point of view of the statistical service to be rendered.

Nevertheless, while this method of assembling internal business facts is generally preferable to either of those discussed above, it is comparatively unsatisfactory. The regular use of records which exist primarily for operating purposes will tend to interfere with the work of both the operating departments and the statistical organization. Under such conditions it is too much to expect that smoothness of cooperation which is indispensable to maximum statistical service. Furthermore, this method necessitates the transcription of a complete list of facts if the basic statistical records are to be adequate to give satisfactory service. Otherwise, the time involved in transcribing the data from records which may at the moment be required for operating purposes, foregoes the possibility of rendering prompt and efficient statistical service. Compared with other methods, this one requires greater manual and mechanical equipment properly to accomplish the work and hence is more expensive.

Statistical Organization: Extra Copies of Operating Records.—In the use of this method the statistical organization assembles in its files, copies of certain primary company records for its exclusive use. This list of records must be sufficiently comprehensive to furnish detailed data such as were outlined in Chapters III, V, and VII. Generally speaking, copies of the records indicated as sources in Chapter X will be adequate.

Copies of these records can generally be obtained with but little additional burden on the operating departments. Usually a carbon copy can be made and forwarded to the statistical organization at the time the record is being prepared for operating use. In certain cases, where it is not feasible to obtain copies of the records or where their infrequent use will not justify it, they may be routed through the statistical organization so that the required data may be transcribed or they may be lent to the statistical organization until the required data can be taken from them. In the latter case the records may be temporarily moved to the statistical organization or the statistical forces may move temporarily to the operating departments to accomplish this work.

This is by far the preferable method to use in assembling internal facts in an operating company, even though a holding company and similar organizations will make free use of reports and statements prepared at regular intervals or on special request. Efficient statistical service requires that the statistical organization be prepared to meet almost any demand for data within a reasonably short period of time and to build up routine reports as required. This can best be done only in case a complete and comprehensive system of basic data is readily at hand. The current maintenance of copies of primary company records exclusively for statistical use accomplishes these results as nothing else can.

This method places a minimum burden on the operating departments and makes possible a maximum statistical service at a minimum cost. It puts the facts of the business in their original detail at the ready disposal of the statistical organization, which permits of the widest range of computations, summarizations and associations of the data. The statistical organization is for the most part self-contained. Its work may generally proceed uninterruptedly without risk of interfering with the operations of the company and without depending on the operating departments for data to meet immediate needs.

Furthermore, the concentration of the statistical work made possible by this method results in the most economical service from the standpoint of personnel, mechanical equipment, supervision, materials and so forth. Likewise, more careful attention can be given to the detailed statistical work and the most serviceable

general reports, special or routine, which involve the facts of more than one department, can be developed more satisfactorily.

Original Collection.—The development of an original study within the company will follow the general method of procedure outlined in Chapter XIV, following. The difficulties encountered will generally be less serious than those involved in a similar study in the external field. As such problems are usually of a special nature and comparatively infrequent in a company where a comprehensive program of basic data is maintained, no further discussion will be offered of this phase of the subject. The series of articles on "A Survey of Delivery Operations" by Harold B. Wess, which appeared in *Administration* from January to June, 1923, will illustrate the original collection of internal data, the first of the series applying specifically to the collection process.

Facts Suggested in the Outline.—Extra copies of operating records will be used generally in collecting the facts outlined in Chapters III, V, and VII. Copies of the records noted in Chapter X will be obtained currently and kept on file in the statistical organization.

Of the *personnel records* listed arrangements can generally be made to receive copies of the following:

Application for Employment
Employes' Individual Record Changes
Pay Roll Form (Dummy)
Report of Absent Employes
Report of Tardy Employes
Separation Form

All the internal personnel facts generally required as basic data can be obtained from these records.

Copies of the *financial records* listed will not be so generally required. Special statements prepared by the treasury department or accounting department either at regular intervals or on request will usually serve. This is particularly true in case of the records listed under Acquisition of Capital. In the other fields copies of the detailed Balance Sheet and Income Statement will usually be

secured each month and copies of the others listed can generally be obtained if the need calls for it.

The same method applies to *operating records,* although the problem is much more complicated. Extra copies of the purchasing, producing and marketing records may be obtained currently with little additional burden on the operating departments. Great care must be exercised, however, to make sure that the forms assembled as basic records cover the situation completely, and that provision be made for taking account of any subsequent adjustments in the principal operating records received. This method may be supplemented in the field of operating facts by establishing certain routine reports to be prepared currently by the several operating departments.

(For references see page 210.)

CHAPTER XIV

THE ORIGINAL COLLECTION

Problems of the Original Collection.—The problems involved in this type of collection are generally different from those involved in the primary and secondary collections. Since no informants are involved, the many questions relative to establishing good faith with the informant, familiarizing him with the inquiry, and stimulating favorable action, are absent. For the most part, the problems arising in the original collection are confined to: (1) the selection of the time, place and conditions of the count, (2) the devising of means for making the count, and (3) the methods of recording the observations.

Time, Place and Conditions of the Count.—It is highly important in the original collection that careful consideration be given to the time of the count, the place of the count, and the conditions under which the observations are made. Generally, we wish to make observations under normal conditions and hence we must see to it that there are no abnormal factors entering in unduly to influence the facts which we wish to record.

If, for example, a street railway company wishes to determine the number of passengers who normally board its cars at a certain street or streets, it would not properly send out enumerators to make a count on a holiday, Sunday, midnight, or any other time when traffic might be considered as abnormal. If, on the other hand, the problem were to secure a measure of traffic on holidays, Sundays, or at midnight, the count would naturally be made at these times. Likewise, in making a count of the automobiles that pass a certain point in considering the location of a gasoline station we should not want to make the count on a day when there was a downpour of rain or when the thoroughfare was piled with snowdrifts. The student should select other original collections and analyze the importance of the time, place, and conditions of making the count in connection with each of them.

Method of Making the Count.—The means of making the count varies so definitely with the type of the enumeration that it is difficult to generalize in regard to this phase of the collection. Furthermore, one or many units may be involved in the count and one or many enumerators may be employed to make it. In some cases also permission may be required from public authorities or from private individuals. These and many other problems arise to complicate the situation.

In any event great care must be taken to develop facilities for making the count which will enable the enumerators to select readily the particular units to be counted and to eliminate others. These things should be completely disposed of before the counting begins: a clear conception of the unit that is to be counted, permission for making the count, and definite methods of procedure.

Furthermore, careful provision should be made for checking the observations. This is particularly important in the original collection, since it is very expensive in time and money to go back and get the correct data when errors have been made. Such checking can be done in different ways as indicated in Appendix E. The best method of checking in this type of collection is through repetition. We can send the enumerators in pairs and, if there is danger of collusion, a supervisor also to see that the counts are strictly independent. Broader checks may be had also from similar data and through common knowledge of the facts in the case.

Methods of Recording the Observations.—The method of recording the facts in the original collection will vary with the character of the investigation and the method adopted for permanently recording and organizing the data. Hence, the method in any given case will be selected primarily according to (1) the convenience to the enumerator and the consequent efficiency and accuracy in the count; (2) convenience in recording the data in the office.

Generally, the data are recorded in one of four different ways, namely: (1) by means of the mechanical counter; (2) by means of the tally sheet; (3) by means of a checking system; (4) by means of a pre-arranged form on which the facts are written in.

The mechanical device is confined primarily to problems in which the data are to be cumulated while only method #4 is used where it is necessary to write out the answer. Further discussion of some of these methods is given in Chapter XXII, beginning on page 217.* The original collection made by R. H. Macy & Company in a study of delivery methods and costs, noted above, will serve as a concrete illustration of this type of collection.

Conclusion.—While most of the data required by a business organization will be obtained through the primary and secondary collection, most concerns at times and some concerns quite frequently secure statistical data through the original collection.

Typical of such collections are the recording of data in time studies in a factory or office; the recording of the calling and called numbers or the holding time in connection with local telephone messages; the counting of pedestrians who pass a certain point on a city street to determine the location of a certain type of retail store such as tobacco, candy, soft-drink, ice cream; the counting of the number of families in a community to determine the advisability of opening a retail store or to get an estimate of the potential demand for a certain product; and so forth. This method is used particularly by advertising agencies and commercial research organizations. It is frequently the only source of the data required and in many other cases it is the most fruitful source.

(For references see page 210.)

CHAPTER XV

THE PRIMARY COLLECTION—GENERAL

Nature of the Primary Collection.—In its practical aspects, the primary collection differs in many ways from the other two. In the primary collection each original observer must be approached individually, while in the original no informant is involved, and in the secondary there is no direct contact with the original observer and the data from several of them may be obtained at one point. Its use likewise involves problems relative to soliciting the informant's interest, familiarizing him with the inquiry, and stimulating his favorable action, all of which are absent in the original and secondary collections.

In the primary collection we may secure data from the informant in a simple request; we may make frequent requests of the same informant; or we may make general arrangements whereby data will be freely given or exchanged at any time.

It is important to note in this connection that the data obtained through the primary collection may be in their original detail or in any sort of derived or computed form. Likewise, the informant may take the data from systematically kept records; he may make a special count for the purpose of the inquiry; or he may furnish the facts from his general knowledge. This conception differs from that sometimes advanced which indicates as an essential characteristic of the primary collection that the data be in their original detail. It will be recalled, however, that the conception here set forth is consistent with the general classification of statistical sources outlined on pages 80 and 81.

Confined Generally to External Field.—The internal data used by the practical statistical organization are collected generally through the primary collection although the original collection is sometimes used. This is true of both the data maintained currently on the basic statistical records and those collected through special studies. The collection of internal facts, however, involves gen-

erally a very special form of the primary collection and is therefore given special treatment in Chapter XIII.

Consequently, this discussion of the methods of collecting data through the primary collection will be developed from the point of view of the gathering of external facts. The primary collection in the external field may be used in the collection of basic data, regularly, or in connection with special studies. The collection of data from other business concerns as outlined in Chapter XIX will perhaps represent the most extensive use of the primary process in collecting data regularly for basic statistical records. This source of external data is so important to the individual business organization that the method of collecting facts from it are given special treatment. Other illustrations may be found in special types of business concerns, such as those interested in the market aspects of agricultural products, which regularly collect facts designed to reflect the crop conditions of the country.

Aside, however, from the collection of facts from other companies, the practical business concern will make comparatively little use of primary collection in maintaining basic statistical records. The use of primary collection will be confined generally to special studies and even here the more general type of business organization will not employ it extensively. Special types of business such as advertising agencies and service organizations will use the primary collection more extensively.

However, the limited use of the primary collection does not detract from the importance of its discussion in a treatise of this kind. The proper examination and evaluation of the data taken from secondary sources requires intimate knowledge of the general principles and methods of the procedure in the collection process. Furthermore, proper classification and analysis of statistical data can be done only with full knowledge and appreciation of the collection process which produced them.

Selection of Informants.—The problem of selecting informants in the primary collection is frequently a very difficult one. It has not been solved when we have merely noted a group of informants who have the data we want. Generally, in practical work, the required data can be obtained from several different

groups of informants. The problem, therefore, usually involves a consideration of all the different groups of informants and a selection from them of the particular group most satisfactory in the specific case.

Facts on the distribution of automobiles, for example, may be obtained from state license bureaus, from manufacturers, from local dealers, or from owners. In undertaking a collection of this type of information the inquirer must decide which group or groups of informants will best serve in the proposed collection. The same is generally true in the collection of all kinds of statistical facts which the statistical organization of a business concern will undertake to get together.

Generally, in making the selection in any given case, the relative characteristics of the different groups of informants with regard to reliability, ease of access, amount of data which can be furnished, and willingness to comply with the request, are important. The significance of these elements in the situation is evident. Each group of informants should be carefully examined with respect to them and the one selected which best serves the purpose in mind.

Methods Used in the Primary Collection.—In general, there are two broad methods of collecting data through the primary collection process: the questionnaire and the personal interview. These two methods will be used separately or in combination, the details varying with the specific case. The questionnaire is fully discussed in Chapter XVI and XVII, and the personal interview in Chapter XVIII.

The Ultimate Ends.—In the primary collection the inquirer should keep in mind the two-fold ultimate end to be accomplished —the greatest relative number of replies from the informants approached, and the highest percentage of "good" replies among the total received. The use of the primary collection can be generally justified only with a relatively large number of replies which are fairly complete and otherwise satisfactory, whether the particular method used be the questionnaire or the personal interview.

The Immediate Ends.—In the accomplishment of this two-fold ultimate end, the inquirer will have certain definite immediate aims in view. In general, the difficulties encountered in a primary collection will be in the nature of the following with respect to the informant:

Indifference
Suspicion
Desire to color the data
Unfamiliarity with the details of the inquiry
Lack of time
Positive opposition to the inquiry which he considers an imposition.

These difficulties must be overcome if the primary collection is to be fruitful. Three specific aims will be set forth in attempting to meet them. Regardless of the method used, the inquirer will so direct the inquiry as to

1. Establish good faith with the informant and thereby allay suspicion and claim his confidence.
2. Familiarize the informant with both the general aspects and detailed features of the inquiry.
3. Stimulate favorable action on the part of the informant.

Establishing Good Faith.—In presenting the inquiry, the informant should be taken into fullest confidence. If the inquirer has a worthy purpose in asking for the data, he should present his case fully and completely—he should place his cards on the table; if not, the inquiry should not be made. The informant should be approached as one whose assistance is required in developing a worthy study. He should be appealed to as working with the inquirer in the accomplishment of a piece of work of interest to them both, the inquirer after all doing the greater part of the work. With such relations the informant should therefore be told all about the proposition.

There are certain definite things which he has a right to know before complying with the request. Who is sponsoring the inquiry? What other informants in the community are being approached? What specific use is to be made of the data? Who will

get to see the data? These and many similar questions come naturally to the mind of the informant and in most cases they must be completely and satisfactorily answered before the best cooperation will be given.

No question of importance should be left unanswered in the informant's mind. The sheer worthiness of the inquiry, fully and completely stated in all its aspects, should be offered to prove the good faith of the inquirer. Extensive use of the primary collection must not depend on sharp salesmanship practice. It must depend on the willingness or cooperation of the informant, acting deliberately when the essentials of the inquiry are at hand. The result of a primary collection in which the confidence of the informants has not been secured is questionable.

Familiarizing the Informant with the Inquiry.—The inquirer should take great care to give the informant full and complete information relative to the inquiry. If he is to furnish data he must know what data are wanted, and, if he is to comply voluntarily with a request, the work involved must be reduced to a minimum. The fact that he is being asked for information presupposes a certain acquaintance with the general field of the inquiry. This must not be too greatly depended on, however, in his interpretation of a specific request. He should be made familiar with the special inquiry presented in both its general and specific aspects. Every phase of the inquiry which the inquirer feels the informant might not readily understand should be carefully and clearly explained. In brief, every step of the inquiry should be so prepared and presented that the informant may not only determine therefrom just what data are wanted but may do so with minimum effort.

The need for such action is two-fold. In the first place, the informant will not ordinarily put forth great effort in order to understand a request, from the compliance with which he derives no tangible benefit. This is especially true if he feels carelessness on the part of the inquirer has added to his burdens. His natural reaction in such cases is either to disregard the request entirely or hurriedly to throw together some material which seems to meet the requirements of the inquiry. On the other hand, if

the inquiry is so presented that the informant can readily understand it, the work involved in answering the inquiry is reduced just so much and the possibilities of a serious effort to comply with the request increased just so much.

In the second place, thorough understanding of the inquiry is absolutely essential to satisfactory results even from informants who seriously undertake to answer the inquiry. The inquirer will secure uniformity, accuracy, and completeness in the detailed data just in proportion as he succeeds in getting all informants to observe the same detailed directions and procedure in submitting the information. For instance, in collecting wage statistics in a primary collection, specific information must be given; otherwise one informant may return wages of male and female employes combined and another wages of only male employes; or one may return earnings and another base rates; or one may include and another exclude part time employes. It is a responsibility of the inquirer to see that the detailed method of procedure is uniform and he can accomplish this only by furnishing complete detailed information to the informants.

The detailed methods of familiarizing the informant with the inquiry will differ according as the questionnaire method and the personal interview method is used. In any event, great care must be taken to see that the informants may readily understand the inquiry and that all may follow the one detailed procedure in preparing the data requested. Complete discussion of the subject in the case of the questionnaire is given on page 177 and in case of personal interview on page 186.

Stimulating Favorable Action.—It is not enough, however, merely to establish good faith with the informant and thoroughly to familiarize him with the inquiry. What the inquirer ultimately desires is a satisfactory reply to the inquiry. Stimulus to favorable action must be added. Generally, such stimulus will not only tend to insure a reply but will also tend to develop a more serious effort on the part of the informant and hence to produce a better quality of reply. Great care must be exercised, however, in the use of any sort of positive stimulation, particularly certain types of personal benefits such as cash payments for replies, lest the inform-

ant, without regard for accuracy, throw together some kind of data or venture careless replies in order to make the data complete and thus secure the benefit.

The use of positive stimulation is based on the broad fundamental business principle that any expenditure in time and money should net a return. This principle is characteristic of human nature. The normal human being looks for the return to himself for any effort put forth. Quite naturally, therefore, the informant, whether an individual or an organization, will extend the fullest cooperation in connection with a request for statistical data only when he is satisfied that he is getting some sort of return, tangible or intangible.

The positive stimulation which the inquirer will use may take a number of different detailed forms. In general, they may be grouped in three broad classes—(1) those netting an immediate personal benefit to the informant, (2) those which challenge the informant's interest in general research and scientific development, and (3) those based on personal or business contacts. The importance of these classes or types will vary with the type of the informant and the nature of the inquiry.

Immediate Personal Benefits.—When the desire for personal benefit is appealed to a favorite practice is to offer the informant a copy of the summarized results of the survey or a copy of the completed study. This is always a method worthy of consideration. The decision as to whether it should be used must be made after careful consideration of the facts of the individual case. Primarily, the decision will be made on the basis of the type of informants. If the inquirer, for example, undertakes the collection of data from the residents of a certain community seeking information as to the extent to which a certain household electrical appliance is used in the community, the individual householder would find little stimulation in the promise to tell her how many persons in her community used the particular appliance. On the other hand, if we should undertake to collect from electric lighting corporations data reflecting the number of residence houses in their respective communities that were electrically wired and should offer to give any individual informant data on a comparable basis

—say, in per cent of total number of families—for other communities, the offer would probably be effective.

This form of stimulation will be more effective as statistical data are used more extensively. Informants will then appreciate that they are getting a return for their efforts in the form of the completed study or report.

Cash payments are sometimes made to the informant for his reply. This is generally inadvisable, as the informant is apt to be more interested in getting the money than in furnishing accurate and complete data.

Personal Benefits Combined With Advertising.—It is often feasible to combine stimulus of this type with advertising and make both more effective. The inquirer may offer the informant an advertising sample of some sort in return for his cooperation in furnishing the data requested. Specifically these stimuli may take a number of different forms, according to the nature of the inquiry and the thing to be advertised. Samples of a commodity in which the inquirer may be interested may be used. Or the informant may be offered a map of particular interest; pertinent statistics of his community, if a business man, such as the number of families and the number in the family; a booklet on the proper management of the informant's particular line of business or how to keep his accounts; and so forth. The local merchant in the smaller city of 10,000 inhabitants or less might be very much interested in a map showing in detail the marketing radius for his line of business and the total amount of retail business of his particular line that is done in the community. All of these things might be put up in the form of advertising and the expense charged to advertising. It is sufficient here to emphasize the effectiveness of the method. The details can be worked out only in connection with each individual case.

Interest in Scientific or Research Developments.—Sometimes the appeal can be made effectively on the basis of the informant's general interest in scientific or research developments. This stimulus cannot be used very extensively. It is confined to those informants who as a general proposition are interested in

the development of the scientific method. The number of individuals or organizations, however, falling in this class are relatively few. While the college president, the professional economist and statistician, or the research expert might be affected by such an appeal, the factory worker, the small retailer, or the average citizen would generally find no appeal in it whatsoever. Usually, we do not greatly enthuse over a vague and remote contribution to the scientific method.

This type of stimulus should be considered and used wherever the type of informant would seem to indicate its effectiveness. It is perhaps the least effective and most rarely used, especially in the field of business.

Personal or Organizational Contacts.—A stimulus which is generally very effective, especially as between organizations, is based on friendly relations—good will. The name of the individual or organization sponsoring the inquiry is the only appeal that is required. The data are given not because the informant expects to get some tangible benefit therefrom but primarily because of the good will of the informant.

The development of the spirit of cooperation between large industrial concerns today bids fair to make possible an extensive use of the primary collection in this field. Competition is giving way to cooperation and on the basis of this cooperation business concerns are coming to exchange with great freedom information even of a confidential nature. The relations thus developing greatly facilitate the collection of data by one business organization from others in the community. Where these relations have been properly developed, to ask for the information is generally assurance that a serious effort will be made to furnish it. In such cases the effectiveness of the primary collection is guaranteed by the congenial business relations existing between the inquiring and informant organizations.

This is the most effective form of stimulus. More detailed discussion of its use in connection with the questionnaire and personal interview is given on pages 166 and 185 respectively.

Specific Application.—The foregoing represents a discussion of the primary collection in its broader aspects. When we under-

take to follow this general procedure in the actual collection of data we find there are certain detailed considerations of importance which are characteristic separately of the questionnaire method in the one case and the personal interview method in the other. These will be developed fully in connection with the discussion of the separate methods in Chapters XVI, XVII, and XVIII.

(For references see page 210.)

CHAPTER XVI

THE PRIMARY COLLECTION—USE OF THE QUESTIONNAIRE

The statistical organization of a business concern in gathering data through the primary collection may resort to the use of the questionnaire, sometimes called also a statistical schedule or question-blank. The extent to which the questionnaire may be used will depend on the character of the inquiry. Generally speaking, it cannot be depended on at the present time for elaborate statistical development. Extensive discussion of the subject must be justified on the basis of its potentialities rather than its actual service at present.

We may here consider, in general, its definition, construction, distribution, and serviceableness.

Meaning of the Questionnaire.—From the practical point of view the statistical questionnaire may be defined as a written request from one person or agent to another for statistical information. The agent who sends the request may be called the inquirer, investigator, or enumerator, and the one to whom it is sent, the informant, or reporter. Throughout these discussions we shall use *inquirer* in the one case and *informant* in the other. The inquirer or the informant may be an individual, or a group of individuals in the form of an organization or association, public or private.

This conception of the questionnaire is a comprehensive one. It includes both the more elaborate type consisting of several pages and the simplest kind of inquiry which is presented to the informant by written communication or even by telegram. Any attempt to establish a finer and more technical definition will carry little of value for the practical statistical organization in a business concern.

Construction of the Questionnaire.—When the inquiry depends entirely on the questionnaire, particularly if the questionnaire is an elaborate one, great care should be taken in its preparation both in regard to the questions and the explanations and directions accompanying them. In the more elaborate questionnaire the wishes of the inquirer are indicated in writing generally by a series of questions prepared in accordance with the principles outlined in Chapter XVII. The questions may be many or few, simple or complicated, according to the character of the inquiry. In any event, the inquirer depends on the questionnaire to state his case and the informant to understand the inquiry; and the data are recorded by the informant without assistance from the inquirer. Detailed discussion of the method of its construction is given in Chapter XVII.

Distribution of the Questionnaire.—When the questionnaire is used exclusively, the inquiry is communicated to the informant in writing. The distribution by mail is most generally used. In special cases, however, where hurried response is required the special messenger in a single community or telegram in more extensive territory may be advisable. The use of the telegram is particularly satisfactory where data are collected regularly from the informant and his readiness to give conscientious response and his familiarity with the data required have been established either through special arrangements by letter or interview or through frequent requests.

Whatever method of distribution may be selected, care should be given to the time of distribution. It is well to release the inquiry so that it will reach the informant at the most propitious time to solicit his favorable reaction to it. If feasible, we should select a time when the informant is not so busy or certainly we should avoid times known to be "rush" for him. If data are requested as of a certain date the inquiry may well reach the informant at about the time the data are available if they are regularly computed, or at an earlier time if special computations are to be made for the inquiry. Furthermore, if it is desirable in the particular case, the questionnaires should be released so as to reach the several informants at about the same time. These several suggestions will serve to illustrate the importance of giving thorough

consideration to the question of the proper time of distributing the questionnaire.

In the distribution of the questionnaire by mail great stress is sometimes placed on the importance of certain details such as the use of first rather than second class mail, the personal rather than form letter, the enclosure of a self-addressed envelop or card for reply, and so forth. There is considerable difference of opinion relative to the importance of these items. Some will contend that the use of the two cent stamp in the distribution will guarantee the attention of the informant and consequently net more satisfactory results while others report practically no difference in the use of the two- and one-cent stamp. Some contend that a warm personal letter makes a stronger appeal while others point out that such a communication smacks of an unjustified cordiality and produces an unwholesome reaction.

We should, of course, take advantage of such of these devices as are fundamentally serviceable. As a matter of fact, however, it is difficult to believe that the collection of data for a substantial and comprehensive statistical development can depend greatly on things of this sort. Beyond observing the usual courtesies in a business communication, little of fundamental value is gained through the introduction of such details in distributing the questionnaire by mail.

A second request by way of follow-up may be distributed if it seems advisable in the particular case. If this is contemplated, the effect of the follow-up should first be determined by a sample distribution to a relatively few representative informants who have not responded to the initial request.

Number of Returns.—In addition to the proper construction of the questionnaire, which is largely within the control of the inquirer and which is discussed at length in Chapter XVII, the relative number of returns in any given case will depend primarily on the following conditions which are generally beyond the control of the inquirer:

1. *The nature and extent of the inquiry*—A simple inquiry calling for only a few facts that are readily at hand, or one in which the informant is interested or which at least avoids asking for

data which the informant considers somewhat personal or confidential, tend to net a higher percentage of returns. On the contrary, a more complicated inquiry which calls for considerable work on the part of the informant, or one that requests facts which the informant generally prefers not to release, will tend to yield a relatively smaller number of returns. Practical inquiries fall at every point within the range of these two extremes and always the nature and extent of the inquiry affects the reaction of the informant and consequently the number of returns.

2. *The type of informant*—The relative number of returns will likewise be affected by the general type of the informant. If the informant is of a type that can more readily understand the inquiry and can appreciate its purport, a higher per cent of returns may be expected. This is true in spite of the fact that the properly constructed questionnaire will conform generally in phraseology and content to the type of informant. Furthermore, if the class of informants is generally accustomed to receiving and replying to questionnaires or to sending them, a higher percentage of returns may be expected. On the other hand, the relative number of returns will not be so large if the informant is of the type, such as a factory worker, that is not accustomed to answering statistical inquiries and is ordinarily not sufficiently educated to understand or appreciate their value.

3. *The relation existing between the inquirer and the informant*—The number of returns to be expected will be greatly affected by the relation existing between the inquirer and the informant. This relationship will vary widely. At the one extreme, the inquirer has authority to compel replies—as in the case of the government agency which is frequently given such power by law, or a corporate business organization in collecting data from its subsidiary, associated or controlled companies. In such case the number of returns is of course very high, frequently approaching 100%. At the other extreme, the inquirer not only has no authority to compel response but does not even have special contact of any sort, as is typified by the collection of data from individuals in the community, for example, in connection with a marketing problem. Here the number of returns depends entirely on the presentation of the case in the questionnaire and the per cent returns are conse-

quently low. Between these extremes we find all kinds of special relations such as those existing between companies who stand in the relation of trader and customer, or, between banks and insurance companies and firms to whom they have lent money, or between business concerns in general where there is a general willingness to cooperate in such matters. Likewise, if requests to the informant have been made in the past with which he has complied or if he has himself made requests of the inquiring organization, he will be more inclined to reply.

It should be emphasized that in using the questionnaire method the presence of such relationship between inquirer and informant is far more effective than any amount of work that may be spent in the construction of the questionnaire itself. It is strongly urged, therefore, that such contacts or relationships be made use of whenever possible. Even in presenting an inquiry to informants with whom you have no special relations it is highly advisable thoroughly to consider the possibility of soliciting the assistance of some agent, individual or organizational, who is influential with the informants being approached. This will tend to establish good faith with the informant and insure recognition of the request.

Frequently, the assistance or endorsement can be obtained from certain organizations who may be interested in the data being collected, such as chambers of commerce, trade associations, boards of trade, and so forth. In such cases, the organization may send a separate letter to the informant soliciting attention to your inquiry or such a letter may accompany the questionnaire.

In spite of the foregoing, it is highly important from the practical point of view that an estimate of the relative number of returns be made beforehand. This is essential for a two-fold reason—first, to enable us to determine the number of questionnaires to distribute in order to secure the desired number of "good" returns, and, second, to enable us to estimate the cost of the collection by the questionnaire method.

It is evident from what has been said above that no definite statement can be made relative to the per cent of returns that may be generally expected in the use of the questionnaire. Efforts to generalize along this line have contributed little of practical

value. While the experience in the use of similar questionnaires by ourselves or others will be of some service, nevertheless, we cannot estimate from generalizations or specific cases the relative number of returns to be expected from a particular questionnaire. In practice, the particular inquirer must estimate what per cent of returns he may expect from his particular questionnaire, distributed to a particular group of informants at a particular time and under particular conditions.

Such an approximation will be arrived at primarily by making a preliminary distribution of a small number of the questionnaires to representative informants and noting the per cent returns from these. The relative number of returns received in similar inquiries conducted by ourselves or others will serve as a broad check. Particularly, if the returns from the preliminary distribution are less than those generally received in similar inquiries, a careful reexamination should be made of every phase of the questionnaire and improvements made wherever possible. Other benefits will be derived from both the preliminary distribution and the inspection of similar inquiries in connection with the construction of the questionnaire.

Quality of Returns.—It is logical to suspect that the quality of the data returned in the questionnaire method is low compared with those obtained through personal interview. Even if we assume a conscientious effort on the part of the informant, the results would still be unsatisfactory in many respects. The inquiry depends entirely on the single communication and regardless of the effort of the inquirer to perfect it and the informant to understand it, errors and incompletions will occur. The inquirer cannot possibly foresee all the problems that will arise in each case and, since there is no immediate supervision of the informant by the inquirer, the disposition of any difficulties which the informant encounters in understanding the inquiry and preparing the reply must be left to his own judgment. It is too much to expect that these decisions will be generally uniform and otherwise satisfactory particularly if the inquiry is at all complicated.

As a matter of fact, however, the assumption of a conscientious effort on the part of the informant is not in accordance with the

facts. Generally speaking, quite the opposite is true except, perhaps, where the question of authority enters in to control the situation. As a result, the informant, tending to escape the burden which the inquiry places on him, hurriedly prepares a reply with only a half-hearted concern for its conformity to the requirements of the inquiry or turns the questionnaire over to a clerk for attention with the instructions that she "do not spend too much time on it." In either event, there is not much hope that the final reply will be a masterpiece of careful work. Furthermore, in practice, there is danger that an improper selection of informants may result from the natural response to the inquiry, for the reason that those who do reply may have been actuated by strong personal feeling on the subject or are appreciative of the value of the inquiry or by some other special consideration. It is possible, therefore, that the data obtained will not be truly representative of the group, certain important elements being either partially or entirely left out of the picture.

The general service from the method is considered on page 170 below. Suffice it at this point to say that in the practical use of the questionnaire many answers are illegible, incomplete, obscure, or run entirely apart from the intent of the inquiry. Accuracy, uniformity and completeness depend on the series of questions and the written explanations and directions accompanying them and no matter how perfect the questions and how full and complete the explanations and directions, errors are bound to creep in.

Comparative Cost.—It is frequently urged that a chief advantage of the questionnaire method is its comparative economy. This contention, however, is not unqualifiedly true. In the case of the more elaborate questionnaire such a claim is frequently exaggerated. When we consider the relatively small number of "good" returns, the additional editing required, the probability of concealed errors and non-comparable features necessitating the closest inspection, the added burden generally placed on the informant tending to cause delay and perhaps additional cost in follow-up —in short, when we consider the total average cost per "good" return, the apparent economy dwindles materially if it does not altogether disappear.

Actually, to secure data from a given number of informants requires the distribution of many times that number of questionnaires. As was stated above, a definite statement cannot be made representing the per cent returns generally obtained. Such a figure will be of little practical value even if estimated with the greatest care. In practice, regardless of what the general fact may be, if the per cent returns for the specific inquiry is the thing of importance and the comparative cost will be estimated from this figure rather than from any general estimate, which may not apply in the particular case.

This approximation will be determined as suggested on page 169. When it has been made the cost may be readily computed. Suppose, for example, that the relative number of returns in the given case is 10% and the inquiry calls for 100 "good" replies. In order to satisfy the inquiry 1,000 questionnaires will be distributed. If the total cost of the preparation and distribution of this 1,000 is $150 including postage, printing, paper, clerical labor, supervision, and so forth, then a single "good" return costs $1.50. A comparison of this with the cost per "good" return through other methods will reveal the relative economy, if any, in the use of the questionnaire method.

As a practical proposition a point in relative number of returns can generally be determined in connection with the use of any given questionnaire below which it ceases to be economical in comparison with the personal interview. If the inquiry lends itself readily to both methods this point should be determined at the outset.

Supplemented by Personal Interview.—In general, the questionnaire is more effective when supplemented with the personal interview, particularly if the inquiry is elaborate. The personal interview may be obtained at the time the inquiry is presented, or during the process of compiling the data or at both times. In any event, the informant is responsible for transcribing the data, the personal interview being used primarily for purposes of familiarizing the informant with the details of the inquiry and of perhaps speeding up the reply. This method is different from that in which the inquirer carries a formal questionnaire in collecting data by

personal interview. In this latter case, the questionnaire is not in reality a method of collecting data but a means of guiding the inquiry and of recording the facts collected by the personal interview method.

It need not be emphasized that the best results with this method will be obtained if the questionnaire is supplemented by personal interview during the process of compilation as well as at the time the inquiry is presented. In other words, the inquirer should keep closely in touch with the informant during the compilation and the informant should be urged to communicate freely with the inquirer in case difficulties are encountered. To the extent to which this practice is observed, the informant will have the benefit of the inquirer's familiarity with the request and the inquirer will be able to some extent currently to check the information as it is being prepared.

This method is generally superior to the use of the questionnaire exclusively. The personal interview tends to overcome many objections to the questionnaire, especially if used freely during the preparation of the data. It not only makes for accuracy, uniformity, and completeness but it also tends to speed up the compilation of the data by keeping the inquiry before the informant. It is not generally feasible, however, where the informants are widely separated although even in such cases it may be advisable to use the personal interview to some extent.

Summary—Service From the Questionnaire.—Generally speaking, we are safe in saying that the use of the questionnaire exclusively in gathering data through primary collection is not to be recommended. The number of "good" returns is relatively so small that the method loses the advantages which are theoretically contained in it. Its practical use in the business world today is generally confined to inquiries which are very simple in character, or those in which the relations between the inquirer and informant are particularly conducive to satisfactory response, or those which involve informants in geographical locations which make the personal interview method impracticable. In general, the per cent of returns is so small and the detailed data are so incomplete and in-

accurate as to render unwise the extensive use of this method in a primary collection of any moment.

This is not to say that the questionnaire should be discarded as a method of collecting statistical information. On the contrary, there is potentially a wide field of service for it, but the technique of its use must be greatly improved and there must be greater appreciation by the general public of the importance of collecting statistical data before its full usefulness can be realized. It is highly important as a method primarily because of its possibilities rather than its actual service at the present time. If the method can be developed so that we may reasonably expect from its use a high percentage of returns, fairly satisfactory in accuracy, uniformity, and completeness, it will be of great service in the collection of data through the primary process. We may reasonably hope for improvements along this line both because of the general tendency of the business world to be more liberal in the distribution of business facts, and because of the more extensive use of statistics in business management and the consequent greater appreciation of the importance of collecting data.

In actual practice we shall attain the most serviceable results from the use of the questionnaire through attention, (1) to methods of distribution particularly with reference to the establishment of some helpful contact or relation with the informant, and (2) to the construction of the questionnaire. With even our best efforts in these respects, however, a great many collections which are carried on by the questionnaire method could be more satisfactorily accomplished by the personal interview method either through individual inquiries or commercial service organizations or advertising agencies and with little or no additional cost.

(For references see page 210.)

CHAPTER XVII

CONSTRUCTION OF THE QUESTIONNAIRE

Aim of the Primary Collection.—The ultimate aim in the primary collection is a two-fold one—viz., (1) a high percentage of returns which are (2) fairly satisfactory in completeness, uniformity, comparability, and accuracy. The immediate aims previously noted are—(1) to catch the interest of the informant, (2) to establish good faith with him, (3) to familiarize him with the inquiry, and (4) to stimulate him to positive action.

Now, if the questionnaire method is selected, it is highly important that extreme care and attention be given to its development. Since, in the use of this method, the questionnaire is generally the only communication between the inquirer and the informant, and since the accomplishment of the entire statistical undertaking rests ultimately on its effectiveness, the detailed preparation of the questionnaire is of vital importance.

The Questionnaire Work Sheets.—Before building the questionnaire, there should be a thorough study of the problem being investigated, with particular attention to the immediate purpose or use of the data. Such reasoning should be directed toward answering the following queries:

1. What use do we wish to make of the data or, more specifically, what are the problems we wish to solve?
2. What data are required in the solution of these problems?
3. How do we wish to record or organize these data?

The answers to these questions will formulate a complete, concise, and comprehensive statement of the problem and will lead directly and clearly to the proper formation of the questionnaire. In other words, the construction of the questionnaire must comprehend two principal considerations—(1) what questions are needed to request the required data? and (2) how can these ques-

tions be formulated and supplemented in order to secure satisfactory results from their use?

This line-up should be supplemented by a proposed form of the record on which the detailed facts are to be entered. This will enable us to make sure that we have made provision in the questionnaire for all the required data.

Thinking the Problem Through.—Before undertaking to fashion the questionnaire we should thoroughly familiarize ourselves with the subject of the inquiry—*we should think the problem through.* This will involve the following:

1. *Reading literature on the subject*—This is important in order that we may become familiar with terminology commonly used by practical men in the field and discover any special points which are apt to cause difficulty if not properly cared for in the construction of the questionnaire.
2. *Criticism by others*—It is also helpful to discuss the subject with experienced men in the practical field with particular reference to special points in the construction of the questionnaire, and, if convenient, perhaps, also to submit to them a copy of the proposed questionnaire for criticism. Discussion with those who have made similar surveys will likewise be fruitful. Information as to the difficulties which they encountered and the particular features of their questionnaires which seemed most effective, should prove beneficial.
3. *Examination of similar questionnaires*—Examinations of similar questionnaires used by others is always advisable. This should cover both their technique and content. Detailed inspection of the explanations, definitions and directions that were given, the character of the questions, the mechanics of the questionnaire, and the results obtained, will usually be fruitful.
4. *Information regarding informants*—We should also familiarize ourselves generally with the important facts related to the informants. What language do they best understand? What are their interests? What are their busy and slack times? What in general is their reaction toward requests for information from them? This and other information which will be apparent in the specific case should be obtained. In this connection, it should be noted that we should try to

put ourselves in the informant's place and sense, if we can, his probable reaction to the inquiry. We should ask ourselves the questions: Could I answer these questions for my own company? Would I conscientiously endeavor to answer them if I should receive such a questionnaire from someone else?

Components of the Questionnaire.—The construction of the statistical questionnaire involves three principal parts:

1. The detailed questions proper.
2. Supplemental material including explanations, definitions, and directions.
3. The mechanical arrangement or organization of the material.

More specifically the questionnaire consists of (1) the *detailed questions* which are the main body of the questionnaire, (2) *means for* getting checks on the replies, (3) *explanatory notes* designed to acquaint the informant with the inquiry, (4) *definitions of terms* designed to familiarize the informant with the data required, and (5) directions to guide the informant in making his reply.

The Questions.—The explanations and directions discussed below tend to allay suspicion, overcome indifference, familiarize the informant generally with the inquiry, and make him willing to answer the questions. The detailed questions themselves, however, are the real essentials of the questionnaire and should be prepared with the greatest care. The inquirer must depend on them finally to produce the data he wishes and the way in which they are fashioned and organized is of vital importance. These questions can be developed only in the light of the facts of the particular situation. There are, however, certain general principles which are more or less applicable in all situations.

1. *Suited to the informant*—In preparing the questions the particular type of informant should be kept clearly in mind. Questions that are satisfactory for presentation to a group of college presidents, for example, would be hopelessly unsatisfactory for distribution to a group af factory workers. Likewise, bankers, retail grocers, factory managers, farmers, and

other groups, each will require questions differing somewhat in their detailed features. The form, wording, and character of the questions must be in accord generally with what the informant is accustomed to if we are to get the most satisfactory results. In preparing the questions for the informant, we must speak his language.

2. *No useless questions*—As the inquirer is asking the informant to spend time in filling out the questionnaire, generally without definite return therefor, common decency requires that he lighten the burden of the informant as much as possible —that he should not ask the informant to do more than the minimum amount of work. To this end he will minimize the data called for and confine the questions rigidly to those which are absolutely essential. Additional questions tend to increase the expense to the inquirer in editing and organizing the data as well as needlessly to increase the burden on the informant.

   Furthermore, useless questions will tend to reduce the number of returns. If the informant gets the impression that the questionnaire has been carelessly thrown together without proper consideration for his own time and expense, he will be inclined to consign it to the waste basket without so much as a serious thought. Even though the inquirer is armed with authority it is not wise for him to impose upon the informant.

   In this connection, it is sometimes suggested that we should ask the informant for only detailed facts—that we should not request computed data. The truth is rather that we should call for the data which will serve our purpose and the furnishing of which will be the least burdensome to the informant. These may be in the form of detailed facts, summations, or computations. The percentage labor turnover of a business concern, for example, would probably be less of a burden than a statement showing the detailed data necessary for its computation.

3. *Questions easily understood*—Further appreciation of the interests of the informant will impose on the inquirer the obligation so to frame his questions that to read them is instantly to understand them. The burdens of the informant should be confined to answering the questions: he should not be asked to bear the additional burden of figuring out

involved or indefinite questions. So far as possible the questions should have only one interpretation and this one should be clearly obvious.

The inquirer prepares the questionnaire but in many ways the informant controls and directs the form of its preparation. The inquirer fits the questionnaire to the informant, and the informant fits the answer to the inquirer. The informant bends the inquirer to his will in the preparation of the questionnaire and then follows his wishes in replying thereto. Time and effort spent in simplifying the questions will be amply compensated for in the accuracy and increased number of returns.

Check Questions.—The detailed questions should be so formulated as to provide checks on such specific replies as seem advisable. We may be on the look-out for incorrect replies because of difficulty in obtaining the data, suspicion of the informant, or some other reason. For all such cases check questions should be provided. These checks should always be provided so far as practicable without adding unnecessary questions.

Position of Explanations, Definitions and Directions.—The exact position of the explanations, definitions, and directions will depend primarily on the amount of space available for them. Other things being equal, the following positions are preferable in the order mentioned:

1. At the beginning of the questionnaire between the heading and the detailed questions.
2. Immediately following the list of detailed questions.
3. On the reverse side of the page or pages containing the questions.
4. On a separate page.

This will apply particularly to the more general items which concern the inquiry as a whole. Those which concern specific questions should, so far as convenient, be placed in the body of the questionnaire adjacent to the questions affected. Primarily, the explanations, definitions, and directions should be placed where they are most likely to be read before the informant undertakes to answer the detailed questions and where they are most con-

venient for inspection. In any event, references should be placed at points in the body of the questionnaire which are likely to be particularly affected by any explanations, definitions, or directions not definitely associated therewith.

Explanatory Notes.—The explanations included in the questionnaire should contain the material designed to induce the informant to undertake to fill out the questionnaire and return it. The detailed points to be covered will vary with the individual inquiry. What has been said in Chapter XV, relative to establishing good faith, familiarizing the informant with the inquiry, and stimulating positive action, will be covered here. Generally these explanations will include the following information:

I. Explanations designed to introduce the inquirer, present the general request, and establish good faith with and win the confidence of the informant. Specifically the following should be included:
   1. The name of the agent who is sponsoring the survey.
   2. The name of the agent who is directing the survey.
   3. The authority for making the survey—*i.e.*, government, private organization, or private individual.
   4. The number of copies of the questionnaire that are being distributed.
   5. The class of informants to whom the questionnaire is being sent.
   6. The geographical territory included in the survey.
   7. The purpose in collecting the data.
   8. The specific use to which the data are to be put.
   9. What the inquirer expects to get out of the survey.
   10. References and attached letters of introduction.

II. Explanations designed to familiarize the informant with the general facts of the inquiry. In addition to contributing to the development of the wholesome attitude on the part of the informant, this information will be generally helpful to him in answering the detailed questions. Included here will be:
   1. A specific, detailed, concrete statement of the problem.
   2. A concrete statement of the limitations or scope of the inquiry.

III. Explanations designed to stimulate the informant to positive

action. Whatever means are resorted to along this line should be incorporated here. The methods to use are discussed in Chapter XV.

Definitions.—The way is not entirely clear, however, even after the willingness of the informant and his familiarity with the general aspects of the inquiry have been assured. The informant must be made familiar with the details of the specific inquiry. Simple, clear-cut, but fully adequate definitions of all terms used in the questionnaire, which may not be clear to the informant, should be worked up and sent along with the inquiry. The number and character of definitions required will vary with the type of informant. Terms which are perfectly clear to one type of informant may be entirely foreign to another. The inquirer should put himself in the informant's place and inspect the terms and definitions from that viewpoint.

Particularly should the statistical unit be clearly defined. What was said in this regard on page 136 of Chapter XII will be doubly important here. The informant is left to himself in the decision of questions that arise in his efforts to answer the inquiry such as the inclusion or exclusion of certain detail. He is wholly dependent on the information he can get from the questionnaire and consequently this information should be full and complete.

Ordinarily, the informant will not take any great pains to ascertain the meaning of a term: rather will he "take a shot" and let the inquirer suffer the consequences of his own carelessness. Furthermore, unless the meaning of the terms which the inquirer has in mind are conveyed clearly to the informant, the latter may fill out the questionnaire with some other meaning in mind and the information be therefore useless to the inquirer. Uniformity of treatment by the informants likewise necessitates that all terms be clearly understood by them. In short the entire survey may be defeated through improper definitions of important terms.

Directions.—Complete directions for actually developing and recording the data called for should be included in the questionnaire. The detail of these directions will vary with the nature of the inquiry and the type of informants. Generally, they will include specific directions on the following points:

1. To whom the reply is to be sent.
2. Time of the reply.
3. The number of copies.
4. The form of the data—*i.e.*, whether in detail, summarized, classified, or computed form such as averages and percentages.
5. The degree of accuracy—The degree of accuracy to which any mathematical computations are to be expressed should be indicated in the directions. It is usually inadvisable in collecting business data by the questionnaire method to attempt a fine degree of accuracy since the unavoidable errors in the data are relatively large. Furthermore, an attempt to get too fine a degree of accuracy is apt to defeat the inquiry.

   For the benefit of the informant, however, we should be very specific in indicating the required accuracy. For example, are the items to be stated to the nearest unit, in 100 units, 10,000 units, 100,000 units, or how? If computations are called for, should the results be to the nearest integer, exact fractions, one, two or three decimals, or what? If to the nearest integer, are fractions of 5 and above to be counted one and all below 5 to be dropped; or all over 5 to be counted one and all at 5 and below to be dropped? All such questions should be decided for the informant in advance.

In fact, specific directions should be given on every point of importance in connection with the questionnaire. If the informant is to make an original count he should be furnished with complete directions therefor. If he is to do extensive clerical work for which the inquirer is willing to pay, this should be stated in the directions.

Mechanics of the Questionnaire.—This phase of the subject has to do primarily with (1) the size, shape, and kind of paper, and (2) the arrangement of the material on the paper. The essential considerations in deciding these two problems are (1) convenience to the inquirer in distributing the questionnaire and in organizing the data from the replies, and (2) convenience to the informant in replying to the inquiry.

The paper to be used may be either in card or sheet form

and will vary in all sorts of ways to suit the needs of the particular case. If the inquiry is brief, a return postal card may be used while if it is elaborate a small pamphlet may be prepared as in some of the government inquiries. The reproductions given in Part 6 of Appendix E will serve as illustrations. Obviously, specific directions cannot be given to apply in detail in any given case. We simply have to use our judgment, bearing in mind the convenience to the inquirer on the one hand and the informant on the other. If we are to use the replies as basic records it may be preferable to have them on a card or on a loose-leaf sheet. As a general thing, they should be of convenient size for the informant and should conform generally to the records in the office of the inquirer. Rarely should the form be larger than the regular letter-size; it is preferable to use more than one sheet.

In arranging the material the questions may be written out, or a tabular form may be prepared in which the data are to be entered, as shown in Examples A and B respectively of Appendix E, Part 6. If the tabular form is used, the principles outlined in Chapter XXVI, in connection with the discussion of tabulation, should be observed in its construction. The form should be spaced and ruled so as easily to be read, using double or heavy lines or grouped items freely. The form may provide that the replies be written in or that the answer be indicated by making a check-mark in the proper place. The use of the latter is comparatively limited in practical work and in many cases where it can be used the greater possibility of inaccuracy and the less opportunity for checking results make its use questionable.

If the question form is used the questions should be systematically grouped in logical order, adequate space should be provided for the answers, and questions not to be answered or spaces not to be used should be clearly indicated.

If the questionnaires are to be used extensively they should be on a good quality of paper and should be reproduced by printing or other durable method.

Conclusion. — The completed questionnaire should be thoroughly examined before it is released. Explanations, definitions, directions, and questions should be subjected to the most

careful criticism by two or more persons. A few copies should then be distributed and a further review made on the basis of this experience.

The original request may be supplemented by one or more follow-ups, the preparation of which will follow along the same lines as that of the original. Usually, no changes will be made in the questionnaire, the letter of transmittal only being adjusted.

The discussion set forth in this chapter applies primarily to the more elaborate type of questionnaire. Simple inquiries will, of course, require less extensive consideration and many of the suggestions offered herein will not be pertinent. No matter how simple the inquiry, however, each point in its construction should be carefully considered before it is released.

(For references see page 210.)

CHAPTER XVIII

THE PERSONAL INTERVIEW

Number of Inquirers.—The personal interview is one of the most effective methods of collecting statistical data in practical work. The inquiry may be completed and the data secured in a single interview if the inquiry is simple and the data requested few and readily at hand. If the data requested are such as to require considerable preparation, the inquiry may be merely presented at the initial interview and the data obtained later, in which case subsequent interviews may or may not be arranged.

In the development of a given inquiry a single inquirer may be sufficient or the particular task to be accomplished may call for the services of a number of inquirers. In general, the primary collections undertaken by the statistical organization of the business concern can be accomplished with few inquirers although in some cases, generally special studies, a large corps may be required. Special types of business organizations such as advertising concerns, magazines, and mail order houses may make extensive use of the primary collection and may employ a large staff of inquirers either permanently or for temporary service in the development of special studies. The number of inquirers, however, will not fundamentally affect the method of procedure.

Qualifications of the Inquirer.—The personal characteristics of the inquirer play a vitally important part in the success of the personal interview method. The government agency may be empowered to compel replies to the inquiry, and attention may be given to the representative of a private organization even though the inquirer himself may not make a good impression, but in case the inquiry is not so supported, the individual inquirer must carry a great deal of the responsibility for the outcome of the inquiry. Furthermore, even where authority or special relations between the inquirer and informant do enter in, it is highly advis-

able to have the free cooperation and good will of the informant. Hence, the inquirer even under such conditions should bring pressure to bear only as a last resort.

The inquirer should possess certain personal qualities which are especially requisite for this type of work. The qualities of courtesy, honesty, and diligence which are generally valuable in any business activity are especially essential here. In addition the inquirer should be extremely tactful, his mind should be delicately sensitive to the reactions of others, and he should possess generally a good personality and manner of presentation. Certain of these qualities are necessary to create and maintain the proper attitude on the part of the informant and others to guarantee that only the facts actually collected will be reported and no fictitious entries be made for the purpose of completing the report; that every reasonable effort will be made to secure the data required; and that the inquirer will not carry any deliberate bias into the inquiry and color the facts accordingly.

Two other most important qualifications are a scientific habit of thought and actual experience in collecting data. The inquirer must be able to recognize at once the facts he wants, to overcome tactfully the evasive, irrelevant, and otherwise unsatisfactory statements of the informant, and to check up on the information given, until he has secured the facts. This can be well done, in the more complicated type of inquiry at least, only by the experienced inquirer.

Instructions to Inquirers.—However well equipped personally the inquirer may be, he must be thoroughly familiar with the inquiry when he approaches the informant. This familiarity will be obtained through careful study and training prior to the actual survey, supplemented, if necessary, by written instructions carefully prepared and furnished to each inquirer to carry with him during the field work. The advisability of preparing written instructions will be decided on the basis of the facts of the particular inquiry. Generally, if the inquiry is elaborate or the number of inquirers large such instructions should be prepared.

These instructions should cover very completely and comprehensively every aspect of the request. In so far as methods of

collection are concerned the inquirer should be thoroughly instructed relative to (1) the method of establishing contact with the informant, (2) the method of presenting the inquiry, and (3) the form of record to use and the method of recording the facts. He should be especially provided with any special inducements which he may use to stimulate action on the part of the informant such as are listed in Chapter XV, pages 157 to 160. In impressing the inquirer with the importance of these things, particular attention should be called to any special conditions or difficulties which may be encountered during the collection process and each inquirer should be urged to report to headquarters particular problems which he may meet in the field work, in order that the proper solutions may be determined and distributed to all inquirers.

Furthermore, if the data required are of such character as to involve bias on the part of the informant, the information should be obtained by indirection. In such case, the specific questions to be used should be carefully prepared and the inquirers should be thoroughly drilled in their use.

The inquirer's preparation for field work will not, however, be confined to the study of methods of collection. He will have to be familiar also with other phases of the inquiry such as the statistical unit, the amount of data desired, the degree of accuracy to be observed, the type of informant, the territory to which his work is to be confined, and the time period during which the data are to be collected. He should also be provided with information relative to the general type of informants covered by the inquiry and, whenever it is feasible, he should have facts regarding the business or other activities of the individual informants he is to meet. Finally, a preliminary survey should be made of a relatively few representative informants and the instructions should be adjusted on the basis of the experience of these interviews.

All of these things should be drilled into the inquirer before he enters on field work. No one should be permitted to undertake the collection of data until he is thoroughly prepared to do so.

Establishing Contact.—A definite method of approach to the informant should be established in advance of the field work and each inquirer made familiar with and instructed to follow it. As in the case of the questionnaire, some special relation between the inquirer or inquiring agent and the informant or informing agent is by far the most effective method of establishing contact. It may be a general letter of introduction, an official request from your own company, a statement from a third party soliciting cooperation for you, a personal introduction, or definite authority as is sometimes present in government inquiries or certain inquiries conducted by business organizations such as the collection of data by a holding company from its subsidiaries. In any event, the time and effort spent in carefully selecting the method of approach to the informant will be amply rewarded. Such a consideration will not only affect the question of obtaining an audience but will also affect the reaction of the informant and consequently the quality and quantity of the results.

While it is not always possible to establish special relations with the informant, nevertheless, that can be done to a surprising extent if proper effort is put forth. In nearly every case in collecting data from an organization special relations of some kind can be established. In collecting data from individuals generally in the community the problem is more difficult, but much benefit can be derived from efforts along the line even in such cases. The inquirer may solicit the cooperation of local organizations such as chambers of commerce, civic associations, women's clubs, churches, schools, and so forth. This cooperation may take the nature of positive and active assistance, particularly if the data collected might likewise be of service to the local organization, or it may be confined to an indorsement of the worthiness of the inquiry. Whatever form the assistance may take it is usually very helpful both in getting the attention of the informant and in securing favorable reaction to the inquiry. Certain other forms of special relations may be serviceable such as the use of subscribers by a periodical or newspaper publication or customers by a retail organization.

In cases where no special relations exist the inquirer will have to depend on his own initiative and ingenuity to establish contact.

He may make an advance appointment for an interview stating only generally the nature of the request; he may present the details of the inquiry by letter asking for an interview and pointing out any benefit which may come to the informant; he may merely announce his anticipated visit at about a certain date without asking a definite appointment; or he may call on the informant without any sort of previous announcement of his coming. In any event, the inquirer should be so familiar with the affairs of the informant that he can time the interview to suit the informant's convenience and he should ever be on the alert to be otherwise courteous and considerate.

Conducting the Interview.—In the interview, the inquirer will generally undertake to establish good faith with the informant, familiarize him with the inquiry, and stimulate favorable action. For the most part, in the use of the personal interview method, the last two of these immediate ends will be accomplished entirely through the interview although the method of establishing contact may also have a bearing on the reaction of the informant. Regardless of the method of contact, however, the satisfactory accomplishment of these ends and hence, the success of the inquiry, will depend to a great extent on the method of conducting the interview.

In presenting the inquiry, the inquirer may be guided by a definite list of questions the specific and accurate answer to each one of which must be obtained from the informant; he may be guided merely by rough notes; or he may depend entirely on his intimate knowledge of the inquiry and conduct the interview without the assistance of any sort of written guide. If the inquiry is at all complicated, it is advisable if not necessary to prepare in advance a written guide to the interview in the form of questions, tabular forms, memoranda or notes in accordance with which the inquiry may be developed. This is even more important if a number of inquirers are involved. Frequently, however, an interview can be conducted more satisfactorily if no written guide of any kind is used. This will be especially feasible where the inquiry is relatively simple and few data are required. Even where a written guide is used the efficient inquirer will soon be-

come so familiar with the inquiry that he can conduct the interview in great part from memory.

Regardless of the presence or absence of a written guide to the inquiry, it is far preferable to direct the interview in a conversational way. This makes the interview a give-and-take proposition, tends to relieve undesirable formality, and is less likely to make the informant feel that he is on the witness stand. It can be done even if a written guide is used but it is most effective where the inquirer merely has a talk with the informant without introducing a formal method of procedure. In most of the primary collections of the business statistical organization the conversational method will be particularly applicable since the informant is generally very much interested in the subject being developed. On the other hand, if the factor of authority enters in, the interview may be conducted formally but even here a conversational form will be quite worth while.

The conversational method requires more time on the part of both the inquirer and the informant and hence the inquirer should be careful not to permit too much time in any one interview. Nevertheless, the added advantage to be gained through the conversational method of conducting the interview will more than offset the increased amount of time required in its use. In this conversational way the inquirer should invite and freely answer all questions relative to the character of the inquiry and the specific use to which the data are to be put until he is satisfied he has the confidence of the informant.

The inquirer should likewise make sure that the informant has a clear understanding of just what is wanted. Any general questions which may occur to the informant relative to the meaning of the inquiry should be readily disposed of and any detailed difficulties which arise in considering the inquiry in relation to the facts desired from the particular informant should be worked out. The inquirer should see to it that the replies offered are consistent with the inquiry. Furthermore, he may assist the informant in determining from the detailed facts at hand just what will be required to meet the conditions of the request. In brief, the inquirer from his point of view and the informant from his, should work together not only to the end that the informant may know in

detail what is wanted but also that what he submits may be consonant with the requirements of the inquiry.

Furthermore, the inquirer can do a great deal to stimulate the informant to favorable action. He can point out the value of the data that are being collected and offer to reciprocate with corresponding data from his own company or to furnish the informant with a summary statement containing adequate information to warrant general conclusions. He may also offer any other inducements in his power as suggested in Chapter XV, page 157, and will tactfully and courteously persist until the informant has given the desired information.

Recording the Data.—In the use of the personal interview method the required data may be prepared (1) by the inquirer, (2) by the informant, or (3) by the inquirer and informant in cooperation.

Data Prepared by Inquirer.—In some instances the informant replies favorably to the inquiry but the work of transcribing the data is done by the inquirer who is given access to the records of the informant from which he takes the data required. To guide him in selecting the facts and to facilitate classification the inquirer may have a formal list of questions or a tabular or statement form or he may transcribe the data in rough form and leave the work of classifying them until later. The particular form to be used will be determined in accordance with the principles indicated in previous chapters and the suggestions given below.

This method of recording the facts is confined of course to the condition where the data are kept on records by the informant and are to be transcribed from these records to satisfy the conditions of the inquiry. It will be further confined generally to cases in which the informant is an organization such as a business concern, a board of education, a chamber of commerce, or board of trade.

The chief use of this method is to be found perhaps among public agencies which frequently resort to it in collecting data from private organizations. It has been used, for example, by the United States Census Bureau and by the Interstate Commerce

Commission, Federal Trade Commission, and state and municipal regulatory bodies in collecting data from business organizations in connection with their regulatory functions. It is not much used by the business statistical organization, although on rare occasions where the work involved is rather extensive it may be advantageous.

This method generally guards against any incompleteness of information which the informant might be inclined to offer or any bias which he might undertake to work in. It also tends to establish uniformity and comparability of data collected from different informants, since the transcription in every case is done under the immediate supervision of the inquirer who is most familiar with the inquiry and most interested in the correctness of the results. Furthermore, it relieves the informant of the work of transcription which is very important if the inquiry is elaborate. It likewise enables the inquirer to place the data on the type of records which will be most serviceable to the statistical organization.

On the other hand, the method is so expensive as to be generally impracticable for a private agency. To furnish the labor forces to transcribe the data from the records of the informants, especially if they are many in number, requires an expenditure usually prohibitive to the ordinary business organization. Furthermore, if the inquiry is such as to require interpretation of the records of the informant or to make it difficult to determine just what detail should be included, the inquirer's unfamiliarity with the records of the informant may lead to errors in the results. Practically, however, this difficulty is usually overcome by the inquirer's consulting freely with the informant.

Perhaps the most serious objection to this method is the disinclination on the part of the informant to have an outside representative or working force handle his official records. In addition to the inconvenience of furnishing space and working facilities, there is the more serious disturbance of having the regular operations of the organization disturbed. The records from which the data are to be transcribed will be required by the informant and he will be generally disinclined to have his own use of them interfered with. On this account it is generally

preferable for the inquirer to pay for having the work of transcription done by the informant. This relieves the inquirer of the burden of doing the work and relieves the informant of the inconvenience of having outsiders handling his records.

For the most part, this method is not serviceable in collecting data through the primary process. It will be used only in very special cases and then the greatest care should be exercised both in planning the work and in actually carrying it out.

Data Prepared by the Informant.—In this case, the inquirer presents the request and leaves the preparation of the facts to the informant. The request may be presented purely by interview, the informant making rough notes, or a memorandum or tabular statement may be left with the informant which outlines the inquiry and gives general instructions and directions relative thereto, to be modified or amplified during the interview to suit the particular situation.

This method differs from the questionnaire supplemented by personal interview in that in it there is no attempt to prepare in advance a complete and comprehensive formal statement which aims to present the inquiry in full and complete form. The details of the inquiry are developed in interview with the informant so as to convey the inquirer's meaning to him and modifications and explanations are made to fit the conditions of the particular informant. The initial interview may be followed by others during the process of compiling the data.

This method is generally preferable to having the inquirer transcribe the data. While it places the burden of compiling and recording the facts entirely on the informant this is preferable to the inconvenience and disturbance of having his official records used by outsiders. On the other hand, it relieves the inquirer of the work of transcribing the data and thus enables him to cover more ground with an equal expenditure of time and money. However, if the inquiry is complicated and the matter is left entirely with the informant after the initial interview, he may delay the preparation of the data or undertake it in only a half hearted way and the information finally secured may be generally unsuited to the purpose for which it is intended. If subsequent interviews

are held during the compilation of data, this method should, in the main, produce satisfactory results.

The prepared data may be delivered to the inquirer by mail or otherwise or he may call for them in person. The latter method affords an opportunity to discuss the data with the informant and thus to accomplish any needed explanations or corrections.

Prepared by the Inquirer and Informant in Cooperation. In many cases the data are recorded during an interview or conference between the inquirer and informant. The factor of cooperation will range from a very low degree in interviews where the informant is very reluctant to furnish the information and the inquirer has to use his every resource to get the facts, to the other extreme where the informant freely gives his time and enthusiastically enters into the discussion of the inquiry and the preparation of the data. The aim will be, of course, to secure the highest degree of cooperation through the method of establishing contact and of conducting the interview emphasized on pages 185 and 186 above.

If the inquirer does not have the free cooperation of the informant the problem of recording the data is more serious. The entire attention of both will be concentrated on the interview, the inquirer bending all his efforts to securing the data and the informant just as deliberately trying to avoid giving them. It is largely a selling job and time taken to record data frequently interferes with the effectiveness of the interview. If the nature of the inquiry and the type of inquirer are such as to make it feasible, it may be advisable under such conditions to make no effort to record the facts during the interview but to trust to memory to note them and then record them later.

In this event the interview may take the form of a conversation by way of merely "talking things over." The informant should be led to ask questions regarding the inquiry or about the inquirer's company. After covering the ground of the inquiry the conversation may be switched to general matters such as general business conditions, or the informant may be led to inquire about the inquiring company. Nothing is more effective than so to win the informant personally that he will want to be courteous

and cooperative. It may be necessary in such cases to make a second call in order to check up on the data. This may not be so bad even though it takes more time, for it furnishes an added opportunity to get acquainted with the informant and win him personally.

On the other hand, if free cooperation of the informant is secured, the request is thoroughly worked over by the inquirer and informant in cooperation in the initial interview and subsequent interviews if necessary. Not only is the general class of facts carefully outlined to the informant but also detailed questions are disposed of as they arise such as the inclusion or exclusion of certain groups of detailed facts, the degree of accuracy to be observed, and so forth. Furthermore, any necessary adjustments or modifications of the data furnished by any particular informant are worked out in these interviews. The inquirer who is thoroughly familiar with the nature and purpose of the inquiry and the informant who is thoroughly familiar with the detailed facts to be furnished work together until the data requested are on a satisfactorily uniform and comparable basis.

The form of recording the data will vary depending on the nature of the inquiry and the wishes of the statistical organization. The facts may be roughly recorded during the interview, the inquirer merely jotting down in irregular form barely enough material to enable him to prepare a finished report later, or a previously prepared tabular form may be used on which the data are recorded systematically. If similar information has been collected from the informant at some previous time, the inquirer may prepare in convenient form the most recent data received from the informant, which may be merely checked in some cases and in others may serve as a guide in the preparation of the data. This is particularly true if a general understanding has been arranged with the informant whereby data may be obtained from time to time or may be regularly exchanged as is frequently the case between business concerns.

In any case, the information may be recorded in the form of rough penciled notes if more convenient. The practice sometimes followed of insisting that the informant approve the data with his signature, even where it is feasible, is not to be recommended.

The inquirer is not nearly so much interested in having the data officially approved as in having them accurate and thoroughly suited to his needs. The cooperative development of the facts by the inquirer and informant is the best guarantee of this result and the signature of the informant is relatively unimportant.

More detailed discussion of the specific types of records to be used and the method of determining the one to use in any specific case is given in Chapter XXII.

This is generally the preferable method of preparing the data if it is feasible in the particular situation. It is free from the burden to the inquirer and inconvenience to the informant, incident to the inquirer's doing the transcription, and it guards against the possible bias and delay involved in having the informant do the work. The fundamental requisites in collecting data are that they be uniform, complete and accurate. This necessitates the uniform and proper disposition of the many questions that arise during the process of preparation or compilation; the inquirer on the one hand and the informant on the other, working in cooperation should produce the highest type of results.

Accuracy, Uniformity, and Completeness.—It is of final importance that the data collected be accurate in every important detail, uniform in character, and thoroughly complete for the purposes at hand. To this end the inquirer should during the interview carefully check up the data offered by the informant. He should evaluate the replies to the main questions of the inquiry and add supplementary questions, if need be, until the information submitted is satisfactory to him from the standpoint of accuracy, uniformity and completeness. He should consequently be well equipped with related questions which will serve to check the replies given. He should likewise check the replies of one informant against those of others in order to catch any outstanding variations, and he should then undertake to find the reason for these variations.

The question of accuracy, uniformity, and completeness, however, extends beyond the data reported by a single inquirer. It involves also the comparison of data reported by different in-

quirers when more than one are used in the inquiry and in such cases additional problems of the same sort arise. Many questions come up during the process of collection which cannot possibly be anticipated and uniformity in the results necessitates uniformity in the disposition of these questions. Several different individuals, however, acting independently, cannot be expected to dispose of them uniformly without special attention to them.

Consequently, as pointed out on page 183 above, the instructions to the inquirers should not only cover specifically all types of difficulties that can be anticipated but they should also emphasize the importance of the inquirer's reporting to headquarters any special problems which he encounters. In this event the proper disposition of the difficulty should be immediately distributed to all inquirers. Likewise, frequent conferences of the inquirers to discuss special problems that have developed during the actual work will be beneficial. If conferences are not feasible as may be the case in a very extensive inquiry, similar results may be obtained by having a member of the statistical organization visit the inquirers in the field. In most primary collections undertaken by the statistical organization of a business concern the inquirers will be in the office daily and consequently the questions raised above can be readily disposed of.

In the personal interview method the fate of the inquiry is largely in the hands of the inquirer. If he is honest, industrious, and able the results should be generally of the highest type. On the other hand, if he is dishonest he may make fictitious entries without even approaching the informant, if he is interested in the inquiry in any biased way he may deliberately color the data, or if he is generally slow intellectually he may accept and report any sort of worthless information that an unsympathetic informant may care to foist on him.

As a matter of practice, inquirers who perfectly satisfy the qualifications cannot always be secured. The practical selection usually takes the form of choosing the best that are available. Some may be very good, others fairly satisfactory, and frequently many that have to be used fall short of the standard in many respects.

Consequently, great care should be taken to provide checks on

the work of the inquirer. The most effective means is to have another inquirer or a representative of the statistical organization make supplemental calls on informants regarding whom reports have been received, particularly those in connection with whom the data reported seems generally incorrect. A second method of checking involves the comparison of the data collected with corresponding information from other sources. This is relatively only a broad check but it serves a very useful purpose. Other methods may be devised to suit the particular case. In any event, the importance of the inquirer's part in the inquiry makes it highly advisable that practical checks such as these supplement the attempt to select the inquirers who fill the qualifications outlined above.

Expensive in Time.—The use of the personal interview method is frequently limited, however, by the time involved in its application. The time can of course be reduced by increasing the number of inquirers but this automatically increases the time required in preparation for the inquiry and furthermore there is a definite limit to the number of satisfactory inquirers who can be assembled. If the informants number a great many and are distributed over a relatively wide geographical territory, to conduct a personal interview with each of them will usually require a great amount of time. Even if the informants are relatively few and confined to a relatively limited locality the personal interview method will be expensive, at least in time, if the interview is elaborate.

Summary—Service from the Method.—It was emphasized on page 155 that the immediate ends to be accomplished in the primary collection were to establish good faith with the informant, to familiarize him with the inquiry, and to stimulate him to favorable action. Generally speaking, the personal interview will accomplish these ends as no other method will. The personal contact of the inquirer with the informant and the opportunity for discussion which the interview offers lend an effectiveness to this method which is distinct. The inquirer is at hand to see that the informant reacts favorably to the request, that he knows what is wanted, and that he may be prevented from submitting

data, intentionally or unintentionally, which are erroneous or otherwise unsuited to the purpose intended.

The personal interview is therefore unquestionably the preferable method in the primary collection. This is particularly true in the case of extensive or involved inquiries which call for a great deal of detailed data. The time and money available from the inquiring agency and the time of the informant are about the only factors which tend to limit the scope of an inquiry by this method. The discussion of the request by the informant and the inquirer provides a means for handling the most complicated type of inquiry. The inquirer from his point of view and the informant from his point of view, working together in cooperation, will produce the highest type of results at the lowest relative cost.

(For references see page 210.)

CHAPTER XIX

COLLECTING FACTS FROM OTHER BUSINESS CONCERNS

Two Methods.—Two broad methods of procedure may be followed in undertaking to assemble facts from other concerns—(1) the statistical organization may depend on trade associations or commercial service organizations for such facts or, (2) the statistical organization may collect the data directly from the other companies. The former is potentially the preferable method but the latter is for the most part the most practicable method at the present time.

Commercial Service Organizations.—The value of a commercial service organization in supplying statistical data is discussed at some length in Chapter XI, page 123. The economy of such a method, where it can be used, is equally evident in the collection of data from other companies. However, there are practical difficulties in this case in the nature of the disinclination of business concerns to release the facts of their business to commercial service organizations. On this account there is faint prospect of being able to depend to any great extent on the commercial service organization for statistical data of other business concerns. Any extensive cooperation of this kind will probably be confined to the trade association, using the term in its broadest sense, or to some specially created organization which is under the control of the business establishments concerned or has their complete confidence.

Trade Associations.—Frequently, the trade association of a particular industrial class will undertake to furnish a statistical service to its member companies. This service may include both internal facts of the several member companies and external facts of general interest to the trade. As regards the internal facts, however, such service is generally limited. In most cases no attempt is made to maintain comprehensive basic records of the

detailed facts of the several member companies. The service usually takes the form of collecting statistical data from members of the association, combining them into statements in which the identity of the individual company is lost, and distributing the combined or summarized statements to the member companies. The data collected are generally routine reports regularly submitted by the companies to the association supplemented at irregular intervals through special requests presented by means of questionnaire or personal interview.

In some cases, the service which the individual company receives from such a practice is excellent and efforts along this line should be given the most liberal encouragement. Full development of this practice would probably lead to the most satisfactory and economical method of collecting data from other companies. If the work now done by such associations were extended and a sufficiently comprehensive program of facts covering the member companies were currently maintained and made available to the individual companies, great economies would be realized. Through the exchange of information by different trade associations the broader picture for the community could be obtained.

At present the services to be derived from such organizations by way of furnishing the facts of other companies is limited. The data available from them cannot take the place of a systematic program of facts such as is outlined in previous chapters. For some time to come, perhaps, the individual company will have to depend on its own efforts for such facts, regardless of the evident wastefulness to the industrial community through great duplication of effort. Further discussion will be confined to the methods whereby the individual company may collect required data directly from other companies.

Data Collected Directly.—While the principles outlined in the previous chapters for directing the primary collection will be generally applicable in collecting data directly from other companies, yet this is so distinctly a special form of the primary collection as to call for the development of special methods.

Much depends on the particular method of collection as to just how complete the data from this type of source may be. While

there has been a marked change in the attitude of business concerns on the question of exchanging information, nevertheless the situation is not all we should like it to be and any plan for collecting data from them should receive the most careful and painstaking consideration. If the wrong method is used and an unwholesome reaction develops therefrom, the problem is made much more difficult and the service from this type of source greatly limited.

Presenting the Inquiry.—Even though there is a general willingness on the part of the larger business concerns to exchange information relative to their respective organizations, a great deal depends on the proper presentation of the inquiry. Generally speaking, it should be presented to some one in the informant organization who has authority to give final answer to the request and who can, if favorably inclined, refer the inquirer to the particular member of the company who can furnish the data, with the direction that the fullest cooperation be extended. The inquiry should of course have the proper official sanction of the inquirer's company.

The personal interview is unquestionably the preferable method of presenting the proposition. A responsible member of the informant organization, usually a general officer of the company, should be approached and the essential features of the inquiry should be fully explained. The inquirer should make his connection through proper officials of his own company and should be provided with a letter of introduction from a general officer thereof when making his initial visit. This first approach to the company is highly important. Future relations and the general question of cooperation depend to a great extent on it. The general suggestions offered on page 185 relative to establishing contact are doubly important here.

Permanent Relations.—The most serviceable method of procedure is to undertake at the initial interview to establish permanent relations with the informant company whereby the facts of the two concerns may be freely exchanged at any time. In this case the proposition presented is not in the form of a specific inquiry but the broader question of establishing the practice of gen-

erally exchanging all types of facts whenever desired. This makes it all the more important that the initial interview be held with a responsible official of the informant company and that the inquiring company be fully authorized to extend the cooperation which is asked.

The establishment of such relations generally means that future contacts with the informant company will be confined for the most part to the junior officials or supervisors who are immediately responsible for the records from which the data are taken. Since they have been generally authorized to cooperate in the matter, the highest grade of service can be obtained by dealing directly with them. The question of authority to release the facts has been disposed of once for all and the task of collecting the data has been reduced to the mere work of transcribing them properly from the records of the informant company.

Regular Collections.—When permanent relations have been established the details of the request are developed with the particular member of the informant company who is to furnish the data. In the initial interview the inquirer and informant will work over the complete list of basic data requested until the facts of the two companies are on a satisfactorily uniform and comparable basis. The inquirer will have with him at these interviews the facts of his own company corresponding to those he is seeking from the informant companies and will be ever ready to secure for them any additional data which may be requested relative to his own organization. In the initial interview the facts of the particular company are adjusted as required to conform to the general scheme of the inquirer. In collecting personnel data, for example, occupational classes are carefully matched up and arranged in comparable groups, methods of wage payments are converted to a comparable basis, any disparity in time worked is adjusted, and other differences eliminated. Likewise, in the field of financial and operations data, the facts are compared in detail and necessary adjustments made.

This comparability must not only be between two concerns but the data of all the concerns included in the survey must be comparable. The inquirer must seek such information from each

individual establishment as will enable him to compare the facts with those of his own concern without materially altering the organization of the facts of his company.

The problem of keeping these basic data current is fairly simple. A representative of the inquiring company will visit each of the informant companies monthly, quarterly, semi-annually or at such other intervals as may be required in the particular case. The frequency will depend generally on the type of data being collected and particularly the rapidity with which the facts change. At these subsequent interviews the inquirer will turn over to the informant the facts of his own company brought up to date. It is also advisable for him to have in convenient form the facts obtained from the informant company at the most recent date which may be checked as being still effective or may be changed as required. This plan will greatly reduce the time required for the interviews.

In either the initial or subsequent interviews the facts may be actually recorded at the time of the interview or they may be worked up later by the informant. In the latter case, the inquirer and informant should keep closely in touch with each other during the preparation of the data.

Subsequent collections indicated above are sometimes made through the use of the questionnaire. In view of the fact that many of the difficulties in using the questionnaire have been removed in the establishment of permanent relations, this method may be satisfactory. The informant is generally willing to co-operate and is also familiar with what is wanted. This is particularly true in the regular collection, where identical detailed facts are regularly required at different periods of time. When it is used, generally a tabular form is furnished on which the data may be entered, or the data of the informant company as of the most recent previous date may be submitted to be checked or corrected to reflect current conditions.

The personal interview method is, however, the preferable one to use. In addition to the fact that business concerns are generally averse to sending the important facts of their business by mail, detailed questions are constantly coming up which can be more satisfactorily disposed of in conference. Furthermore,

the personal interview tends to establish a more or less permanent and substantial source for data of this class. Its beneficial feature is the fact that the inquirer and informant in cooperation develop the facts required. This not only permits the immediate and proper disposition of any questions that may arise, but establishes in time a personal connection and understanding between the inquirer and the informant which is of great value in this type of collection. Contrary to first impression, such a program is generally not expensive compared with the results obtained. One man of medium salary can take care of the entire program, meet all interviews, and still accomplish considerable work in the office.

Special Collections.—Supplementing the regular collections miscellaneous inquiries may be made of the informant companies from time to time. In view of the fact that permanent relations have been established and personal contacts made with a representative of the informant company, these requests will be easily handled. If the data are simple and readily available in the informant company a telephone call or a letter will suffice. Otherwise a special visit should be made to the informant company and the request fully developed in personal interview.

Form of the Data.—In developing the systematic plan outlined above, the facts should be accepted in the form which is most convenient for the informant company and which still fulfills requirements of the inquiring company. Thus, the data may be obtained from one company in one form and from another company in another form. Care must be taken, however, to make sure they are comparable and will suit the purpose of the inquirer.

Generally speaking, the basic data will impose a minimum burden on the informant company and render maximum service to the inquirer. They can be readily combined with other figures and can be developed into whatever form the statistical service should call for. In some cases, however, computed or derived data will be satisfactory to the inquirer and more convenient to the informant, such as averages in rates of pay, period of service, and production costs; percentages in labor turnover; and percentage changes in rates of pay and labor costs.

(For references see page 210.)

CHAPTER XX

THE SECONDARY COLLECTION

Nature of Secondary Collection.—It is clear that the methods to be used in the secondary collection will be different from those required in the original and primary collections. Instead of having to draw out the facts from an uninterested or reluctant informant as is usually the case in the primary collection or make detailed observations as in the original collection, in the secondary collection it is necessary merely to provide methods for making available data which have been generally released in some form of publication. Furthermore, the inquirer can generally depend on the agency who issues the publication for liberal cooperation. Difficulties sometimes arise in determining just what the data represent but these problems are not so difficult as those of the original or primary collection.

The methods used in the secondary collection vary in detail with the kinds of sources. We shall therefore consider separately the methods generally serviceable to the statistical organization in collecting data from (1) government agencies, (2) non-government agencies, and (3) general publications. For the most part, these sources will be secondary but with respect to some data they may be primary. Likewise, while the statistical organization will obtain data from these sources mostly through the secondary collection, the primary collection will be used in some cases.

Government Agencies.—The methods of collecting data from the different classes of government agencies are so similar that the federal, state, and municipal sources will not be distinguished. For the most part, the data collected by government agencies are published and consequently the practical statistical organization will generally depend on the official publications in collecting facts from this type of source. Our discussion will therefore be con-

fined for the most part to methods of collecting from government publications.

It may be advisable occasionally to depend on a commercial service organization for data from this kind of source. The serviceableness of such organizations has been discussed at some length in the previous chapter. Generally speaking, however, the practical statistical organization will find it preferable to make arrangements to secure copies of advance releases in regard to selected data, to keep at hand copies of certain official publications, and to inspect additional publications for data required in special studies. It may also be advisable in special cases to keep closely in touch with the government agencies through personal interview.

Special Inquiry.—In rare cases special requests may be made for additional facts or special combinations of facts which are not available from the publications. The methods of directing such requests will, therefore, be generally similar to those used in the primary collection as outlined in previous chapters. Wherever feasible the personal interview will be preferable, particularly if the inquiry is at all complicated. Otherwise, written request may be made, in which case the principles suggested in Chapters XVI and XVII should be observed.

Advance Release.—There will be certain data from government sources which the practical statistical organization will want currently and at the earliest possible time after their collection and compilation. The specific types of such data will vary with the requirements of the particular statistical organization. Some will be interested in price data or cost of living data such as are published by the U. S. Bureau of Labor Statistics, others in crop estimates developed by the U. S. Department of Agriculture, still others in wage data such as are worked up regularly by various state governments, and others in other classes of data.

For all such facts, the statistical organization should arrange to receive advance information whenever possible. In some cases advance information is released before the formal publication as a regular practice and in others special arrangements may be made to have certain facts forwarded in advance. Specific arrangements will be developed in the light of the requirements of the particular

company on the one hand and the facilities of the government agency on the other.

Copies of Publications.—Certain statistical publications issued by government agencies will be generally serviceable to the statistical organization. Illustrative of such publications are the *Monthly Review* of the U. S. Bureau of Labor Statistics; the *Federal Reserve Bulletin,* published by the Federal Reserve Board; the *Statistical Abstract of the United States, Commerce Yearbook,* and *Monthly Survey of Current Business,* issued by the U. S. Department of Commerce; *Income Statistics,* published by the U. S. Treasury Department, selected volumes of the federal decennial census, and of the federal census of manufactures. The statistical organization should make a survey of both the regular and special publications of such government agencies as may seem advisable and by inspection select those which should be currently at hand. The statistical organization should, by means of the agency's regular mailing list or otherwise, arrange to keep in touch with its statistical activities and in this way copies of required publications can be secured as issued.

Inspection and Research.—There are many data collected and published by government agencies whose infrequent use by the statistical organization will not justify keeping currently on file the publications which contain them. Such of these data as may be required generally in the case of special studies will be collected through research inspection of the publications which are commonly available in libraries. In undertaking the collection, therefore, the source of data required will be determined as definitely as possible, through the means suggested in Chapter XI, on pages 114 to 119. This will serve as a general guide in the research work. Further methods of procedure will follow along the lines suggested below in connection with the collection of data from general publications.

Non-Governmental Agencies.—For the most part, the methods outlined above under "Governmental Agencies" will serve for collecting data from non-governmental agencies. Educational Associations will be handled almost exclusively through their

publications, which may be either kept currently at hand or examined as required. Commercial Service Organizations and University Bureaus will generally be covered through subscriptions whereby regular publications and special reports are received and, in some cases, special investigations undertaken upon request. It may be noted that with respect to industrial organizations and associations and the miscellaneous agencies, advance releases are not so generally available as in connection with government bureaus. In collecting facts from publications of non-governmental agencies, the same general procedure will be followed as is outlined below in connection with the collection of data from general publications.

General Publications.—A great deal of information of interest to the management of a business appears in general publications. To make provision whereby pertinent data from this field may be made systematically available for the service of the company is a problem more difficult than appears on first thought. The volume of statistical material contained in such publications is so great and the type and form of publications so varied, that a comprehensive inspection in this field involves not only a tremendous amount of actual work but also constant and careful watching to make sure that all important facts for the purposes of the company are obtained.

To accomplish this end one of two general methods may be used—(1) the statistical organization may depend on various commercial services which may undertake for a certain compensation comprehensively to examine this general class of literature and to furnish the company with certain types of data currently appearing therein, or (2) the statistical organization may undertake to make such examination within its own office and to transcribe the data as required.

Data Furnished by Commercial Service Organizations. The first of these two methods is by far the simpler and perhaps not more expensive. This method consists simply in entering into agreement with the agency either by regular subscription whereby the information furnished is determined for the most part by the agency, or by special arrangement whereby the agency will take

the responsibility of furnishing currently to the company all important data of certain types which appear in general publications such as books, periodicals, technical journals, newspapers, and pamphlets. A list of such organizations is given in Chapter XI, page 124.

The regular subscription is frequently used and much service can be derived therefrom. It can hardly be said, however, that there is any service organization whose work is so distinctly statistical and so comprehensive in scope as to cover satisfactorily any general group of data. The regular service generally includes non-statistical information, the replies to special inquiries, and the willingness to make special investigations upon request either with or without additional compensation. There is generally no announced intention to make a comprehensive inspection of general publications with respect to certain types of data. The service is not specialized to that extent. Consequently, the practical statistical organization cannot well depend on such an arrangement to keep in touch with statistical facts appearing in general publications.

However, if commercial service organizations are available which will adequately supply statistical information appearing in any class of current literature it is usually economical to employ them. They partake of the nature of cooperative organizations, serving many customers, and the expense, which would be largely the same regardless of the number of customers, is distributed to all customers and hence is relatively low to any one.

Service Organization to Furnish Statistical Data.—There is ample room and a great need for a service organization whose primary purpose would be to accumulate and distribute statistical data. Such an organization might take a number of different forms with respect to its detailed character and the scope of its activities. The service might be confined to furnishing basic data or facts organized in simple classifications, associations or computations. Or, it might be extended to include the development of complete statistical studies on request and to keep certain figures partaking of the nature of standards, worked up currently.

In any event, a comprehensive system of basic data would be always at hand, in the field of personnel, finances, and operations

covering both the internal and external aspects of the practical business organization. The entire list of internal facts outlined in Chapters III, V, and VII, would be kept currently at hand so far as practicable for the community as a whole, for certain selected industries, for industrial classes, and other groups, as the need required.

These data would then be combined and classified to suit the specific needs of any subscriber. Out of such facts might readily be established the general rate of pay for particular occupational classes in any selected community, and standard practices regarding vacation allowances, hours of work, lunch periods, and so forth. Similar services would be offered in the fields of financial and operations facts wherever pertinent. Likewise, in the field of external facts detailed data would be kept currently at hand covering all of the important types of data. In this case the actual data might be furnished to the subscriber or only the specific references to them. The detailed character of such an organization would be governed by the demand in the particular case and the commercial aspects of the service.

The business world is waiting for just such an organization and the need for it will increase as time goes on. The assembling of facts is a very difficult task and to a great extent the work of doing this is multiplied as many times as there are different companies collecting similar data. With a service organization, a staff of experts in the collection of facts could be developed who would accomplish the task more thoroughly than the individual business could hope to do. This one effort would then serve the entire business community. Such an organization would not confine its services to providing data from current literature but would be equipped to furnish all types of available statistical data from any source, including private business organizations, government agencies, and non-governmental agencies discussed above. It would be a general fact collecting and fact distributing organization.

The practical aspects of such an effort would have to be thoroughly considered. It might be that a trade association, a university bureau, a chamber of commerce, a library, or some such agency should serve as the controlling body; or some nationwide

organization might be worked up whereby branches could be established as required. Our aim, however, is merely to suggest briefly the general type of such a service and the economies which the industrial community might realize therefrom. Its detailed form could be developed only after extensive consideration. With the present disinclination of business concerns to release the facts of their companies to commercial service organizations, it is probable that such an organization would have to be developed round a trade association, chamber of commerce, or university bureau. At any rate, business concerns would have to be assured of the good faith and unquestioned integrity of the organization before they would cooperate.

Collection Made by the Statistical Organization.—The primary method of collecting data from literature is inspection and research. If this is undertaken within the statistical organization, then by the methods outlined in Chapter XI, page 127, in the discussion of statistical sources, the general classes of literature to be examined will be determined. When this has been done, means will be set up within the statistical organization carefully to examine the publications currently and to make the pertinent detailed facts available for the needs of the company along the lines suggested in Chapter XXII in the discussion of the organization of business facts.

The literature to be examined may be made available by purchase through regular subscription or otherwise or by using copies in public libraries. The former may be expensive and the latter is generally cumbersome and otherwise unsatisfactory. Generally speaking, it is much preferable to provide copies in the statistical organization, particularly copies of those publications which are used extensively. Care should be taken, however, not to overload the organization with periodicals regularly subscribed for. This matter should be very rigidly controlled and only those subscriptions maintained which surely prove their worth. For all special studies, the library facilities are generally quite adequate.

The facts will be generally made available for use in the statistical organization (1) by keeping copies of the publications currently at hand, (2) by keeping systematic record of the references

to detailed facts and transcribing the facts from library copies as required, or (3) by transcribing the actual data at the time the inspection of the literature is made. If the need for the data is such as to require that they be readily available or that secondary basic data be worked up, the detailed facts will have to be transcribed. This will be true even though copies of the publications are at hand. If the use of the data is confined generally to special studies it may be sufficient either to keep a detailed reference list and transcribe the facts from library copies of the publications as required or to secure copies of publications for the statistical organization and by clipping or otherwise keep them on file as basic statistical records.

Detailed discussion of the method of recording facts taken from such sources is given in Chapter XXII.

References—Chapters XII to XX (Collection)

Bailey and Cummings. Statistics, 1920, Ch. III.

Bowley, A. L. Elements of Statistics, 2nd Ed., 1902, Ch. III. (Illustrations).

Hall, S. R. Handbook of Sales Management, 1924, Ch. V.

Jerome, Harry. Statistical Method, 1924, Ch. XVI.

King, W. I. Elements of Statistical Method, 1912, Chs. V, VI and VII.

Rugg, H. O. Statistical Methods Applied in Education, 1917, pp. 39-55.

Secrist, Horace. An Introduction to Statistical Methods, 1917, Chs. II, III, and IV.

U. S. Bureau of Labor Statistics. Bulletin No. 326—Methods of Procuring and Computing Statistical Information.

White, Percival. Market Analysis, 1921, pp. 20-42.

Part II—Organizing Basic Records

CHAPTER XXI

EDITING THE INITIAL PRIMARY BASIC RECORDS

Importance of Editing.—In developing the subject of recording data, in Chapter XXII, the records on which the basic data come to the statistical organization are classified as initial primary basic records. Since the systematic statistical service rests entirely on these records, it is clear that they are of vital importance in the statistical service and must be thoroughly complete and reliable.

In order that these records may fulfill the requirements of a foundation for the statistical service, they must be carefully edited immediately on their arrival at the statistical organization. The work involved in editing will vary generally with the kind of data and the method of collection. Internal data will usually require less editing than external, and data collected by the questionnaire method will for the most part require more editing than those collected by personal interview. Whatever the kind of data or the method of collection, these records should be carefully inspected and given such editing as may be required.

It should be borne in mind that in no case should the original entries on the primary basic records be destroyed. Wherever corrections are made they should be noted separately and distinctly from the original entries and should be clearly designated as corrections. This affords the opportunity always to inspect the original entries and to review the corrections.

In editing the initial primary basic records the aim should be to perfect the records in whatever respect they may be faulty. Primarily, this editing will take the form of an examination for completeness, uniformity, and accuracy.

Completeness of the Records.—Editing the initial primary basic records for completeness consists primarily in seeing that

either the data required in the collection are entered on the record or that an explanation is given for their not being so entered. This point should be kept clearly in mind also in collecting statistical data. Wherever an entry is deliberately omitted during the collection process, this fact should be clearly indicated and the reason therefor given.

First of all, care should be taken to see that all entries on the record are clear both as to legibility and meaning—so far as possible all ambiguous, obscure, and illegible entries should be eliminated. This may involve nothing more than making careful inspection of the entry. It may involve comparing it with related entries on the record. If field inquirers were used in the collection, the particular one who made or secured the record may need to examine it. Or, as a final resort, a supplemental effort may be made in the field to correct the entry. The extent to which efforts should be made to correct unsatisfactory entries is developed below. In any event, clear and definite entries should be made to replace obscure, illegible, or ambiguous ones before the record is finally released for file. Otherwise, the same difficulty will be met whenever the record is used and probably different dispositions made of questionable entries at different times. If the entry cannot be satisfactorily corrected, it should be made a "no report" or "unknown" entry.

Furthermore, care should be taken to see that some sort of entry appears at every point on the record where an entry is called for. In such cases, the proper entry may be deduced from others on the record or we may go to the field for it. If in any given case the record is finally filed without certain entries being at hand, these entries should be covered with "no report" or "unknown."

Uniformity of the Records.—The importance of uniformity in collecting data has been emphasized on previous pages. As there stated, lack of uniformity in any group of facts will tend completely to invalidate them for statistical purposes. We cannot, therefore, be satisfied with our efforts to maintain uniformity during the collection; for whatever these efforts may have been, irregularities are apt to creep in. It is very essential that the records

received in the statistical organization be thoroughly examined from this point of view before they are finally filed.

This is particularly true in the external field where standardization is less general than in the internal field. There will be great variation, for example, among business concerns generally in accounting classifications, occupational classifications, methods of computing labor turnover, the determination of what constitutes absence or tardiness, hours of work, basis of compensation, and so forth. These variations may take the form of the use of different terminology to designate identical classes of data or the inclusion of different detailed facts in classes designated by identical terms. Stenographers and typists may be recorded as "clerks" by one company and classified separately by another; the basis for an absence in one company may be an hour and in another a half day or even a day; a lateness of three minutes may constitute a tardiness case in one company while a half hour may be required in another.

All such variations should be ironed out through careful editing of the records. This is a much more difficult task in many ways than editing for completeness. The lack of uniformity is not so evident on the face of the record. Outstanding variations in the data usually indicate probable lack of uniformity. Since irregularity may not be evident in some cases, however, until the data have been classified, simple classifications or computations should be made for the purpose of testing the uniformity of the entries whenever it seems advisable. Sometimes the examination of related entries will also be helpful in checking for uniformity.

The principal method of adjusting entries which are not uniform is to repeat the collection effort. Generally, it is impossible to make the correction otherwise, although it may be done sometimes from related facts.

Accuracy of the Records.—We cannot be too careful relative to the accuracy of the primary basic data. Each individual record should be put to a most severe test in this regard. We should thoroughly investigate everything which has even the appearance of error. The methods of checking outlined in Appendix E, Part 5, should be used wherever they are applicable.

Certain types of errors will be readily noted but it will be difficult to detect others. In the former group are, (1) errors in computations, (2) certain cases of misplaced entries, and (3) glaring inconsistencies. Computations which appear on the record can be checked and any errors which exist may be generally detected. In most cases, these errors can be corrected although sometimes it may be impossible to determine just where the error lies. In an addition, for example, it may not be apparent whether the total is incorrect or whether the error is in one or more of the individual items of which the total is composed. Certain wrong entries may be easily detected by inspection such as "night worker" in the space provided for age, or "25" in the space which calls for the name of the occupational class. We are frequently able to correct such errors by transferring the entries.

Likewise glaring inconsistencies may be at once apparent on the inspection of the record. If an employe, for example, is classed as a typist and is listed as receiving a compensation of $4,000 a year, or if a factory's capacity is indicated as 1,000 units of output per month and then for a certain month is stated to have produced 10,000 units, or if an employe is listed as being 25 years of age and as having a period of service of 20 years—in such cases it is apparent at once that errors have been made. It is not so easy, however, to correct errors of this kind and generally we have to go back to the source.

A second group which is very difficult to detect consists of minor errors which do not stand out prominently. They may be wrong entries or inconsistencies as indicated above or they may be simply false individual entries. The reversal of the entries for age and rate of pay of an employe who is 30 years old and whose weekly wage is $25 would hardly be detected. Likewise, the error resulting from listing a typist's pay of $1,500 a year as $1,000 will generally not be apparent. Furthermore, the great number of possible errors in the individual entries are generally beyond the closest inspection of the editor, such as an average rate of pay of $20 a week recorded as $22, sales of $100,000 recorded at $110,000, production of 10,000 units recorded as 8,000 and so forth. For the great bulk of the entries we have to depend on the accuracy of the collection. It is very important

therefore that every effort be made to guarantee the accuracy of the data during the collection.

Estimating the Significance of Errors or Incompletions in the Records.—If errors which are detected in the records tend generally to be compensating, they are not nearly so serious as they otherwise might be. In view of the fact, however, that these records are to be filed as basic records and used for all sorts of purposes the same degree of dependence cannot be placed on the compensating factor as can be done in special studies where the significance of the errors can be evaluated in the light of the specific purpose to be served. It is generally very important that the primary basic data of the statistical organization be accurate in every detail.

Similarly, in the matter of completeness of the records, it should be emphasized that if the data at hand in a given case constitute relatively a large portion of the total required and are overwhelmingly of one class, the data which are lacking are comparatively unimportant. If our records contain data showing that 95% of the employes are day workers, it will be of little value so far as this classification is concerned to seek information relative to the other 5%. Even if they were all night workers they would not materially affect the general picture. Likewise, if data at hand cover 90% of the sales of the company and indicate definitely that the company distributes 85% of its goods to a particular class of customers, data on the other 10% are quite unimportant in connection with this particular aspect of sales distribution. In both of these cases if the missing data were all different in their essential aspects from those at hand the conclusions would not be greatly affected. It is a fair presumption, however, that the character of these data will correspond generally to that of the data which are at hand. Nevertheless, in view of the fact that the purposes for which the basic data are to be used cannot be foreseen, in this case also it will be better to have them as full and complete as is feasible.

Conclusion.—The primary burden of this chapter is that the basic data records should not lack in completeness and reliability. They constitute a vital part of the statistical program and any

shortcomings in them will tend seriously to hamper the service. The foregoing discussion which emphasizes the importance of editing them for completeness, uniformity, and accuracy has been designed to give the broader impression that these records should be as nearly perfect as it is feasible to make them. The importance of this point will be emphasized further in the next chapter, particularly in connection with the initial primary basic records.

References

Bailey and Cummings. Statistics, 1920, Ch. IV.
Jerome, Harry. Statistical Method, 1924. pp. 303-304.
King, W. I. Elements of Statistical Method, 1912, Sec. 38, p. 62.
Secrist, Horace. An Introduction to Statistical Methods, 1917, pp. 55-57.

CHAPTER XXII

TYPES OF BASIC RECORDS

Methods of Maintaining Basic Data.—The data required in a program of statistics in a business concern will be provided in two ways—(1) by the current maintenance of a system of basic data, and (2) through special collections. The type of detailed facts generally included in the basic data is illustrated in Chapters III to VIII, and the discussion of records in this chapter will be confined primarily to those required in connection with the basic data.

The Importance of Proper Basic Records.—This step in the development of a statistical service is a very important one. Even though a complete and comprehensive list of detailed facts may have been collected as the foundation of the systematic program of statistics, the statistical service nevertheless will be limited and unsatisfactory unless these facts are conveniently at hand. The data must be so recorded in the statistical organization that their analysis and interpretation may be accomplished with a minimum expenditure of time and money. This is particularly true in connection with special studies as opposed to routine reports and in the proper type of statistical service special studies play a very important part. In the preparation of both routine reports and special studies, however, great economy will be realized in having the basic data recorded in a convenient form.

Primary Basic Records.—Basic statistical records may be classified according to the form of the data into two broad groups—(1) primary basic records, and (2) secondary basic records.

The primary basic records of the statistical organization are the records provided to carry the systematic basic data. They are the foundation of the statistical service and should consequently contain adequate detailed facts covering every aspect of the busi-

ness organization. If the statistical service is to be satisfactory, the primary basic records must be complete and comprehensive. These records will furnish the detailed data—the raw material—from which the great bulk of the completed statistical products are to be developed whether they be routine reports or special studies.

On the basis of their convenience in practical work, the primary basic records may be classified into two types—(1) the *initial* primary basic records, and (2) the *ultimate* primary basic records.

Initial Primary Basic Records.—The initial primary basic records are those on which the data come to the statistical organization. The basic data suggested in Chapters III to VIII, collected from the several sources discussed in Chapters IX to XI and by the various methods developed in Chapters XII to XX, are fed more or less regularly into the statistical organization. They may come from a number of different sources and be collected in a variety of ways. They are nevertheless the basic data initially received by the statistical organization, and the records containing them are therefore classed as initial primary basic records.

All other records as well as routine reports and special statements in connection with the regular service are based finally on the initial primary basic records. They mark the limitations of the systematic service both as to comprehensiveness and accuracy of data. The regular statistical service will for the most part be confined to analyses which can be made from the data available from these records, and the accuracy of the material in any development cannot be checked farther back than these records without going outside of the statistical organization. These records are consequently fundamental and their proper development is quite as important as getting together a comprehensive list of facts.

In this connection it should be emphasized that convenience in organizing the data in the office should be kept clearly in mind in selecting the form of record to be used in collecting the data. Of course, the primary aim in any type of collection is to get the facts and nothing should be permitted seriously to interfere therewith. On the other hand, every effort consistent with the effectiveness of the collection should be made to have the records used in the collection conform to the type most convenient in the or-

ganization of the data. This will generally net a substantial economy in the office work both in time and money. The foregoing, of course, implies that the most serviceable type of record from the standpoint of the organization of the data will have been determined before the collection is begun.

Form of Initial Primary Basic Records.—It must be fully appreciated, however, that the records used in collecting data can be controlled by the statistical organization to a limited extent only. These records are closely identified with the methods of collection and frequently the methods of collection will be the determining factor in selecting the form of record. The relative control which the statistical organization may have will vary generally with the source of data and the method of collecting them.

In the collection of *internal data,* the initial primary basic records may be in the form of primary company records, in which case the statistical organization will have no control over them except perhaps as it may suggest revisions of the company records to make them more serviceable for statistical purposes; or they may be in the form of special statements furnished as a matter of routine by the operating departments, in which case the statistical organization can control the form of record to a great extent. In either case, these records should be in a form as convenient for statistical work as is feasible.

In the *external field* the form will vary with the type of collection. In the original collection the form of the record is largely in the control of the statistical organization and the particular record to be used in the field—hand-counting machine, tally form, checking form, or writing form—can usually be chosen with full consideration for its serviceableness in the office. In the primary collection, the form of the questionnaire and the form used by the interviewer should both be made to conform to the system of records in the office in so far as this can be done without detracting from the effectiveness of the collection. This does not guarantee, however, that all of the replies and reports will be on these forms since replies to questionnaires are frequently received on separate statements, or memoranda, and since it will be preferable many times for the inquirer merely to make rough notes and

transcribe them later or to accept copies of special statements from the informant. Nevertheless, in this type of collection the bulk of the initial primary basic records can generally be made to serve the convenience of the statistical organization in size and form. This is particularly true in collecting data from other business concerns.

In the secondary collection, a copy of the publication, a clipping therefrom, or a transcription of the required data, may serve as the initial primary basic record. Generally speaking, the statistical organization will determine which of these will be used, and consequently the selection can be made to serve the convenience of the organization.

In any event, the fundamental aim in the selection of these records is effectiveness in the collection as has been noted above, and we should not sacrifice effectiveness of collection for convenience in recording the data.

Ultimate Primary Basic Records.—The service of initial primary basic records, however, is generally limited by their form. In most cases it is relatively inconvenient and burdensome to develop statistical studies directly from them. The basic data of the statistical organization are collected from all sorts of sources and in all sorts of ways, and are reported to the statistical organization on all sorts of forms in spite of our efforts to control these forms. Some of the forms will be convenient in the organization of the data and others will be very inconvenient. It is, therefore, frequently advisable to transfer the basic data from the initial primary records to more convenient types of records. These more convenient types we shall designate as *ultimate primary basic records.*

The detailed forms of these types will vary, and no attempt will be made here to offer a comprehensive list. From the broad point of view they may be classified generally in three groups—(1) specially prepared forms for long-hand work, (2) patented card record systems, such as the "Acme," "Costmeter," "Findex," etc., and (3) the Hollerith or similar tabulating systems.

Long-Hand Forms.—The *first* type depends for the most part on the ingenuity of the statistician and will take many dif-

ferent forms. From the initial primary basic records the statistical organization will devise the particular forms which will best serve the needs of the specific case. These records may be in the form of cards or of loose-leaf sheets to be kept in binders, the latter being preferable in most cases. Figure 17 below will serve

WAGE STATISTICS—OUTSIDE INDUSTRIES

DISTRICT—1 CLASS OF INDUSTRY: BANKS, SEX—F CODE—X

OCCUPATIONAL CLASS	JANUARY 1, 1925				JULY 1, 1925				JANUARY 1, 1926			
	No.	Start	Ave.	Max.	No.	Start	Ave.	Max.	No.	Start	Ave.	Max.
Supervisors	5	—	40.50	50.00								
Stenographers	20	16.00	23.00	30.00								
Typists	40	12.00	15.00	18.00								
Office Girls	10	8.00	9.00	11.00								
Clerks—Grade 1	25	10.00	12.00	15.00								
" 2	75	11.00	13.00	17.00								
" 3	100	12.00	15.00	20.00								
" 4	40	16.00	21.00	26.00								
Compt. Operators	7	15.00	20.00	25.00								
Total Empl. In Company	Male — Female — Total —				Male — Female — Total —				Male — Female — Total —			
REPRESENTATIVE INTERVIEWED—*Mr. J. Smith* DATE OF INTERVIEW—12-20-24												

FIGURE 17.—Long-hand Form used as the Ultimate Primary Basic Record for Wage Statistics of Other Companies in the Community

Note that the name of the company is recorded by code instead of by name. A separate form will be used for each concern and by using both sides it can be made to serve for entries at a number of different dates. A 5 x 8 card will generally be sufficient for six entries or a period of three years if the survey is made semi-annually. This form may also serve to record the combined figures for separate industrial groups such as banks, insurance companies, metal trades, or textiles

as an illustration. Additional suggestions will be found in Appendix B.

The method of working up the data for these forms will vary with the character of the original detail, the form of the data required on the records, and the preference of the statistician. Generally he will use—(1) the *tally method* as illustrated in

Figure 18, (2) the *checking method* as illustrated in Figure 19, or (3) *separate cards* as illustrated in Figure 20. Methods of developing classifications, summations, and computations are discussed in Chapter XXXIII, Analysis—General, and following chapters.

The long-hand type of form will be confined generally to the smaller concerns where the relatively small amount of work to

LABOR TURNOVER STATISTICS

Classification of Losses by Rates of Pay and Reasons for Leaving

Weekly Rate of Pay	RESIGNATIONS						DISMISSALS			OTHER LOSSES			Total
	Wages	Work'g C'dn's	Other Emp't	Ill Health	Domes. Re'ns	Other	Poor Atten.	Poor Work	Other	Pen's'd	Dec's'd	Other	
Under 16													
16	𝍸		///		4		/	////					14
17	//	//	𝍸	/	//		//				/		15
18			/										1
19	/				/								2
20		/			/	/							3
21			/	//			/		/			/	6
22	//						/						3
23													0
24													0
25			/		///								4
26				//	/	//							5
27									/		/		2
28	//		/										3
29													0
30				///		/			/				5
31													0
32													0
33			///		𝍸	/							9
34 & Over						/				//	/		4
Total	12	3	15	8	14	6	5	4	3	2	3	1	76

FIGURE 18.—The "Tally Method" used in Classifying Labor Losses by Rates of Pay and Reasons for Leaving

Data are taken from the record on which the loss is reported and entered by tally marks on this form

be done does not justify the use of the Hollerith system. It consists merely in putting the initial primary basic records in a more convenient form by the long-hand process. In general, economy is realized in two ways. First, the number of individual records to be handled when working with the detail has been materially reduced; and, second, the material has been organized into a more systematic form which facilitates the work of tabulation either in the development of routine reports or special studies.

Patent Card Record Systems.—The *second* type of record noted above may be any one of the various patented card records such as "Findex," "Kardex," "Rand," "Acme," "Costmeter," and so forth. This record will likewise be generally confined to the smaller concerns. It is more flexible than the long-hand records noted above and consequently serves well to record data in which

LABOR TURNOVER STATISTICS
Classification of Losses by Reasons for Leaving

Employee's No.	RESIGNATIONS						DISMISSALS			OTHER LOSSES		
	Wages	Work'g C'd'n's	Other Emp't	Ill Health	Domes. Re'ns	Other	Poor Attend.	Poor Work	Other	Pen's'd	Dec's'd	Other
37	X											
51				X								
62			X									
112											X	
114								X				
121					X							
137	X											
139			X									
155		X										
161	X											
167					X							
182									X			
193							X					
195			X									
198						X						
212		X										
215	X											
317							X					
423			X									
567								X				
612							X					
Total	4	2	4	1	2	1	3	2	1	0	1	0

FIGURE 19.—The "Check Method" used to Classify Labor Losses by Reasons for Leaving

Data are taken from the record on which the loss is reported and checked on this form.

the changes are frequent, such as rates of pay, occupational classes, and period of service in the personnel field. Likewise, they are convenient for reference and serve well to carry data which need to be referred to frequently. These systems are very serviceable and they should be given careful consideration in any given case.

The Hollerith System.—The *third* type of ultimate primary basic records suggested is fully discussed in Chapter XLV on Mechanical Aids. The detailed method of procedure in the

SEPARATION RECORD

(Forward to Payroll Department)

_____ COMPANY _____ DEPARTMENT _____ DIVISION

Clock No._____ Pay due to_____ Supervisor_____ Date_____

NAME	POSITION	Date Effective	Rate of Pay Per *Week*	Date Engaged	Employe's Rating (1)	REASON FOR LEAVING	
						General (2)	Specific (3)

(1) Excellent, average or poor.
(2) Resigned, transferred, laid off, or discharged.
(3) Specific reason for resigning, or for being transferred, laid off or discharged. Use standard terms listed on reverse of this form.

REMARKS:

Signed by..... Approved by.....

FIGURE 20.—Record on which Labor Losses are Reported

The losses may be classified by reason for leaving, rate of pay, age, period of service, and so forth, simply by sorting these records into groups

use of this system is there outlined and Figures 116 and 117 will illustrate the form of the record.

This is the most satisfactory method where the volume of work to be done is sufficient to warrant the use of the machines from the standpoint of expense. Hollerith machines—the punching, verifying, sorting, and tabulating machines—are expensive and their use is economical only in case they can be kept busy practically all the time. A tremendous amount of work can be done with them and an organization has to have a large volume of statistical work to keep the machines sufficiently busy on statistics alone. Wherever practicable, however, this is the preferable form in which to have the primary basic data.

The chief advantage of the tabulating cards as a form of primary record is in the speed with which the detail can be worked up. There is no other form of record which will yield the desired result with such satisfaction. There are two disadvantages: first, the one mentioned above, that the expense involved may be relatively heavy in cases where the amount of work to be done is not large; and second that preparing the cards calls for a great deal of work. The problem of expense which applies in a relatively small business concern, may be frequently overcome by utilizing the machines also for things other than statistical data. The amount of work involved in preparing the cards is not a serious disadvantage in any case where the detailed material is used a great deal—*i.e.*, is organized in a number of different ways and at frequent intervals. The transferring of the data to the tabulating cards is generally much less work than taking them directly from the initial records. Furthermore, to take details from the tabulating cards is so much more economical from the standpoint of time and money that there is no comparison with the method of taking them from the initial records. Therefore, wherever the detail is to be resorted to frequently or the volume of work is heavy, tabulating cards are a great economy as a form of primary basic record and make possible the most extensive and satisfactory service from the material.

Forms Only Suggestive.—These are but general suggestions in connection with the different forms into which the primary

basic data may be converted for purposes of convenience. These forms have been listed as being perhaps in most general use and as suggestive of the advantages to be gained through the conversion of the primary basic data into more convenient forms than those in which they come to the statistical organization.

Secondary Basic Records.—The primary basic records contain in systematic form the individual facts required in the statistical service. An inquiry can be answered from them only by developing required associations or computations from the original detail. Frequently such a practice would be expensive both in time and money and could handicap the service by the delay involved. This is generally true even though the system of primary basic records is excellent and the department well equipped with the most efficient and up-to-date mechanical aids. On the other hand, it is not possible to foresee all the forms in which these detailed facts are to be organized. Provision must be made whereby the detail will be available in the shortest possible time and in almost any conceivable form, and every device possible must be resorted to to facilitate the work of furnishing the basic data to the management in the form required.

To avoid the expense and delay of going back to the primary basic records on the one hand, and the unjustifiable expense of completely developing all the basic data into an inexcusably large number of finished statements on the other, a system of "halfway" records should be developed which we shall designate as *secondary basic records*. These records contain certain basic computations or associations whose importance and frequent use require that they be readily at hand but whose nature is such as not to call for general distribution in the company.

They are generally not so detailed as the primary basic records. The units of the primary basic records have been computed or associated to furnish a wide field of basic information which is used extensively in the company. Ordinarily the individual unit facts cannot be identified from the secondary basic records except, perhaps, where development is confined to an association of two or more classes of data. Whenever it is necessary to inspect the individual units to locate an error, either initially or after

the study has been distributed, the secondary basic records will generally not serve; the primary basic records must be resorted to.

Detail to be Included in Secondary Basic Records.—These records are designed to make readily available such basic computations or associations as do not justify routine distribution but whose demand cannot be adequately met directly from the primary basic data. We must decide in any given case what material would better be prepared in the form of secondary basic data and what may be satisfactory in the form of primary basic data. The more convenient and serviceable the form of primary basic data, the less will be the need for secondary basic data.

In general, the secondary basic data should comprise, in connection with each fundamental process, (1) such portion of the primary basic data as have not been developed for routine reports; and (2) greater detail and different combinations in connection with the primary basic data which have been developed into routine reports. The requirements along these two lines will be determined from a general consideration at the outset of what the needs may be, supplemented as time goes on by conclusions based on experience.

The fact of fundamental importance is that adequate secondary basic data must be developed to make the statistical service efficient. If the statistical organization is to serve its purpose well, it must be prepared to furnish *what material* is wanted at the *time* it is wanted. A comprehensive system of primary and secondary basic data will be necessary to guarantee this result.

Forms of Secondary Basic Records.—What has been said on pages 220 to 226 above, relative to the types of ultimate primary basic records, will apply generally to secondary basic records. The use of the Hollerith system, however, requires special note. It may involve the organization of a distinct set of cards for the secondary basic data; or, if convenient, these data may be added to the cards which contain the primary basic data. In the latter case the primary basic data would be recorded in certain spaces on the cards and the secondary basic data in other spaces. For example, in recording the number of employes by periods of service, the actual period of service might be recorded

in a certain space on the card and a summarized classification might be recorded in another space such as, for example, under 3 months, 3 under 6, 6 under 9, and so forth. From one set of facts we could get any classification of periods of service that the basic detail would permit; from the other, the detailed facts summarized in simple groups.

WAGE STATISTICS

Number of Employes by Wage Groups and Periods of Service
January 1, 1925

ACCOUNTING DEPARTMENT CLERKS FEMALE

Wage Group	Total in Group	PERIOD OF SERVICE GROUP						
		0–6 Mos.	6–12 Mos.	1–2 Yrs.	2–3 Yrs.	3–4 Yrs.	4–5 Yrs.	5 Yrs. and Over
Under $10	6	5	1	—	—	—	—	—
$10–12	30	20	10	—	—	—	—	—
12–14	35	16	15	4	—	—	—	—
14–16	18	5	8	3	2	—	—	—
16–18	41	10	10	7	12	2	—	—
18–20	69	5	4	16	27	16	1	—
20–22	64	1	4	7	14	28	8	2
22–24	68	—	—	15	20	17	11	5
24–26	74	—	—	—	6	32	21	15
26–28	22	—	—	—	1	5	12	4
28–30	21	—	—	—	—	—	—	21
30–32	6	—	—	—	—	—	—	6
32–34	6	—	—	—	1	—	2	3
34–36	13	—	—	1	—	2	—	10
36–38	6	—	1	—	—	—	1	4
38–40	9	—	—	—	1	—	3	5
40–50	18	1	—	—	—	—	—	17
$50 and over	—	—	—	—	—	—	—	—
Total	506	63	53	53	84	102	59	82

FIGURE 21.—A Typical Secondary Basic Record in the Field of Personnel Statistics

The detailed wage and period-of-service groups entered are merely for illustrative purposes. These facts will be worked up monthly, quarterly, semi-annually, or annually as required. Other associations might likewise be made such as age, nationality, health classes, with each other or with wages or period of service.

The specific forms to be used in the development of secondary basic data and the specific character of the data themselves will vary greatly. Certain specimen statements are given in Figures 21 and 22 and others in Appendix B, Suggested Reports, which may serve to suggest the general type of forms and data to be maintained on this type of record. They will facilitate the selection of secondary basic records in the practical situation.

PRODUCTION — SALES — DELIVERIES

In Units of Product

MONTH OF_____

ITEMS	MONDAY			TUESDAY			WEDNESDAY			THURSDAY			FRIDAY			SATURDAY		
	Shop Deliv.	Sales Orders	Ship-ments	Shop Deliv.	Sales Orders	Ship-ments	Shop Deliv.	Sales Orders	Ship-ments	Shop Deliv.	Sales Orders	Ship-ments	Shop Deliv.	Sales Orders	Ship-ments	Shop Deliv.	Sales Orders	Ship-ments

FIGURE 22.—A Typical Secondary Basic Record in the Field of Operating Statistics

Routine Reports and Special Studies as Records.—These are the primary methods of distributing statistical information. In time, a great amount of data of the derived or computed form will be available from the several copies of these reports. These reports therefore become in turn very serviceable records of data in the statistical organization. They are of course, unsystematic from the standpoint of records—particularly the special studies—and consequently cannot be greatly depended on. They should always be carefully filed, however, and made use of to the fullest extent. Typical routine reports are suggested in Appendix B.

Classification and Association of Facts.—The importance of the classification and association of facts is emphasized in Chapted XXXIII. It is apparent that this point must be fully appreciated and kept clearly in mind during the process of developing basic records, since classification and association of the data cannot be carried further than the detail shown on the basic records. These records, therefore, should not only give the detailed facts but should carry these facts in such way that all important associations may be readily effected.

References

Bailey and Cummings. Statistics, 1920, Ch. V.
Jerome, Harry. Statistical Method, 1924, pp. 24-28.
Rugg, H. O. Statistical Methods Applied to Education, 1917, pp. 60-66.
White, Percival. Market Analysis, 1921, Ch. III.

Part III—Presenting Statistical Data

CHAPTER XXIII

PRESENTING STATISTICAL DATA—GENERAL

Relation of Presentation to Analysis, Interpretation and Distribution.—The development of a systematic program of statistics in a business organization may be considered as consisting of two broad phases. The first covers the work of making available in the statistical organization the basic data which are to serve as the foundation of the statistical service. The second includes the analysis, interpretation, and distribution of these data. Thus far in these discussions attention has been confined to the first of these two phases of the subject.

From this point forward attention will be given chiefly to the second phase. It will be developed in two parts: (1) the *analysis* of statistical data in which will be discussed averages, frequency analysis, index numbers, and statistical relationships; (2) the interpretation of statistical data.

Running through these three aspects of the subject as an important supplement thereto is the subject of presentation and since it is not associated exclusively with any one of these processes but is important in connection with them all, it will be taken up specifically at this point.

Method of Treatment.—The general method of treatment will be changed at this point. In the discussion of statistical sources and the collection of statistical data the method followed was not only to give a systematic discussion of statistical sources and methods of collection generally but also to point out specifically the sources of the detailed facts required in a program of statistics as outlined in Chapters III to VIII and the detailed methods to be used in collecting them.

Obviously, such a procedure cannot be followed in considering presentation, analysis, and interpretation. Separate treatment

of the three fields of data from these points of view would require many times the space available. Instead, the separate methods of presentation, analysis, and interpretation will be developed, noting in connection with each its use generally in the fields of personnel, financial, and operations facts, with particular emphasis on the special service which it renders.

The Meaning of Presentation.—It should be observed that in the present discussion presentation will be considered as a distinct process, entirely separate and apart from analysis and interpretation. It is very serviceable in the accomplishment of both analysis and interpretation and its chief function is perhaps along this line. It may, however, be used for purposes other than these, some of which may indeed be very important.

More specifically, presentation in statistics involves merely the physical representation of statistical data. It is not concerned with the kind of data, the amount of data, the form, nor the purpose for which they are to be used, although all of these will affect the method of presentation. Questions relative to the kind, amount, and form of the data and their purpose will have been disposed of before presentation is begun. Should the data be presented in the form of averages, percentages, index numbers? Should certain data be included in the presentation or excluded? If computed or derived figures such as averages and percentages are to be used, should the absolute numbers be presented also? All such questions are questions not of presentation but of analysis and as such will be disposed of generally before the question of presentation arises. The problem of presentation begins only after such questions have been disposed of and the specific data which are to be presented have been finally determined. How should the data be set forth to serve best the purpose in mind? This and this only is a question of presentation. Convenience of presentation will be kept in mind during the preparation of the data and difficulties of presentation may lead to an adjustment of the data that were intended for presentation. Generally, however, problems involved in determining what data are to be presented do not fall within the scope of presentation.

It is evident therefore, that the presentation of the data does

not affect them in any way. The data after presentation are identical in every respect with the data just before presentation. They are simply arranged in more serviceable fashion from the standpoint of the purpose in mind. If they are in the form of averages or percentages or totals when presentation begins they will be in the form of averages or percentages or totals after presentation has been completed. Likewise, presentation neither eliminates any portion of the data nor directs that others be included. The question frequently raised in connection with presentation as to whether absolute numbers should be presented along with percentages or averages is decided before presentation is considered. It is not a question of presentation. Presentation takes either the percentages or absolute numbers or both as may have been decided by processes other than presentation and sets them forth in the most convenient and serviceable arrangement for the specific purpose.

Methods of Presentation.—Fundamentally, methods of presenting statistical data may be grouped into three broad classes—textual, tabular, and graphic. The general nature of these classes is apparent from the terms themselves. Their detailed characteristics will be brought out in the more comprehensive discussion of them given in Chapters XXIV to XXXII.

The Purpose to be Served.—As in nearly every phase of statistical development, the purpose to be served plays a most important part in the presentation of data. Generally speaking, each step is selected only after it is given most thorough consideration in the light of the purpose for which the data are to be used. The very same data may be presented in entirely different ways to serve different purposes.

From the broad point of view, there are three fundamental purposes or classes of purposes for which data are presented—(1) to make them readily available for further use, (2) to facilitate analysis and interpretation, and (3) to accompany an analytical or interpretive texual discussion as support to its contents or as reference for the reader. Generally speaking, each broad class or type of purpose has a distinct bearing on the form of presentation to be used. For example, the first type of purpose

indicated above will generally be served by the simple tabular form; while the other types will usually call for the more highly developed tables and graphs. More detailed discussion of the importance of the purpose in presentation is given below in connection with the development of the textual, tabular, and graphic methods.

Presentation and the Selection of Data.—In selecting the data for presentation in any given case the kind and amount of data, their form, and the classes in which they are to be grouped will be determined primarily by analysis. Through classifications, computations, and other analytical processes the specific facts and classes of facts which are required to serve the purpose in mind will be arrived at. Since, however, the problems of presentation and analysis are so closely associated certain special considerations should be noted at this point.

It is sometimes advisable, for example, to include relatively unimportant data in the presentation in order to make the picture complete. In the presentation of age statistics of employes, for example, all ages above 65 or even above 55 may be unimportant for the purpose in mind and yet it may be desirable to present them in order that the presentation may reflect the total number of employes of the company. Similarly, relatively unimportant data may be included in wage, period of service, price, sales, and other kinds of statistics. Such data will be included generally only when their omission may tend to weaken confidence in the main body of data. When they are given, they can generally be grouped in broad classes as is suggested in the discussion of the "classification of data" below.

Likewise, additional data may be required to aid in the proper interpretation of the principal facts. Suppose, for example, that we are making a comparative study of the average rates of pay of typists and stenographers as among the several departments of the company. In presenting this information facts showing the number of employes in each department should be associated with the wage data. The question as to whether there are 50 stenographers or 2 stenographers in one department, for example, will have a distinct bearing on the significance of the average

weekly rate of pay for stenographers in that department. In making a finer analysis the age and period of service of the different groups might also be required. In any given case, whatever facts are required to accomplish the end in view should be included, even though they have only an indirect bearing on the situation.

Form of Data.—Furthermore, when using averages or percentages or other derived figures it is frequently advisable to present the absolute figures also. This will be particularly true in the informative type of presentation or where it is likely that the computations may need to be checked up. This will not be the usual thing in the practical statistical work of a business concern, as the work sheets of the computations are readily available for inspection in the statistical organization. Here, again, the purpose to be served will be the determining factor.

On the other hand, unimportant adjustments can sometimes be made in the data which will greatly facilitate their presentation and will not detract materially from their general effectiveness or service from the analytical point of view.

If, for example, a decided advantage in presentation is gained by doing so, the "miscellaneous" class or classes may be omitted altogether. In such case, the reason for omitting them should be given in explanatory notes together with a full and complete statement of just what they consist of and their bearing on the purpose to be served in the presentation. The classifications may likewise be adjusted, certain classes being eliminated or others added, so long as the integrity of the classes included as most effective for accomplishing the end in view is maintained. Of course the effectiveness of the data should not be sacrificed for convenience of presentation but wherever convenience of presentation can be gained without such sacrifice, full advantage should be taken of the opportunity.

Titles of Statistical Tables and Graphs.—The purpose of the title assigned to a statistical table or graph is to reveal to the reader quickly and clearly the chief elements which the graph or table contains. Two primary conditions govern the selection of the title in practical statistical work—(1) limitations of space, and (2) the necessity for clearness. The first of these requires

that the title be brief; the second requires that it be sufficiently elaborate to reveal the essential characteristics of the contents. The necessity for crowding the title into a small space calls for effort and skill.

In order to have the finished title brief, concise, and very explicit, so that it permits of only one interpretation, it should generally be in the form of a series of phrases rather than complete sentences.

The following items in the order listed are typical of the make-up of a title:

1. Name of company
2. Kind and form of data
3. Organizational or territorial limitations
4. Main classifications and such subclassifications as may be required in the particular case, listed in order of their importance
5. Time period represented by the data

In addition, the title should emphasize such particular elements in the contents of the table or graph as may be of particular importance. The following title for a table or graph on wage statistics will help to clarify this suggestion:

A. B. C. Company
Average Rates of Pay—Departments and Divisions
by Occupational Classes and Period of Service Groups
as of December 31, 19—

As a practical method of procedure in selecting a title we should list the various items which might possibly be included in the title. With this list before us we can decide most satisfactorily what is finally to be included and the way it should be arranged.

The foregoing is offered merely by way of suggestion to aid the beginner. In actual practice we find titles which vary in many ways from the suggested form, both as to arrangement and contents, and which serve satisfactorily. Titles to the tables and figures reproduced in this volume will be illustrative. However, the formulation of a proper title is generally tedious and it

will be helpful to establish some more or less definite method of procedure in its development.

Explanatory and Qualifying Notes.—Explanatory notes are distinguished from footnotes in that they refer to matters which apply to the presentation generally, while footnotes apply for the most part to specific points or items of the presentation. It should be observed that explanatory notes will not be required so extensively in connection with the statistical service of a business organization where full and complete information relative to the data and the processes of working them up is readily at hand in the statistical organization. This is particularly true in the development and distribution of internal facts. However, extreme care should be taken in every case to offer explanatory notes where the conditions require them.

It will be impossible to indicate definitely what should be included in the explanatory notes. Generally, they will include any information which is considered important in reading, using, or interpreting the data. Typical of the items to be included are:

1. *Sources of the Data*—It is usually unnecessary to include sources of the data in the regular statistical service, particularly in connection with internal data. In the external field the case is somewhat different. References to sources should be particularly clear and specific. If the reference is to a publication, for example, care should be taken to note the volume, date of publication, the chapter, page, etc.
2. *Processes of Compiling the Data*—Any part or phase of the process of compiling the data which is important in connection with their use or interpretation should be included. If the average is used, for instance, the type should be indicated, *i.e.*, whether it is the mean, the median, the mode, the geometric or what not; or if estimates are used the method of estimation should be outlined.
3. *Limitations of the Data*—It is especially important that the limitations of the data be clearly indicated. This may involve an explanation of certain aspects of the method of collection, together with a statement relative to the statistical unit, and to the degree of accuracy observed in the collection and compilation of the data. If, for example, in the collection of

absence statistics, absence cases are considered only to the nearest half-day, this fact should be noted in the explanatory notes lest some one assume that they are only to the nearest day or, on the other hand, to the nearest hour. Likewise, if in the compilation of data decimals have been disregarded, the fact should be noted.

4. *General*—In addition to the foregoing, anything else necessary to the proper interpretation and use of the data and especially anything necessary to prevent incorrect deductions from or improper uses of them, should be included in the explanatory notes. This is particularly true in the case of a special study. Any limitations of the original data or of those being presented should be carefully noted. The explanatory notes should fully cover whatever is necessary to give the reader a full and complete understanding of the original data and of any computations, combinations, or associations.

Footnotes.—Footnotes are designed to clarify specific points in the table or graph. They should be used freely in connection with any specific entry or group of entries which are not clear from the presentation itself. The notes accompanying the data of Table 1 of Chapter XXV will illustrate. It is obvious that the entries in this table to which the footnotes apply would be entirely misleading without them.

Selecting the Method of Presentation.—The method of presentation to be used in any given case will depend on the character of the data and the purpose to be served. If the data are a few concentrated figures, they will probably be woven into the textual discussion. A mass of detailed items to be used primarily for reference would demand tabular presentation. If the purpose is to facilitate analysis or interpretation either the tabular or the graphic method may be used; or the two methods may be used together, the graphic supplementing the tabular or the table furnished to support the graphic. The tabular method saves space and clarifies the data in comparison with the textual method; and the graphic method is designed to clarify and emphasize the essential features of the data which are not brought out clearly in the tabular or textual methods. We can therefore select the method of presentation only in the light of the facts of the particular case.

In this connection it should be strongly emphasized that the graphic method, which is very expensive, should be used only where it contributes something definite. There is a general tendency today to throw everything into graphic form but frequently the graph adds little of value. Graphic devices should be used freely when they clarify; but if they do not clarify, they should be dispensed with.

Furthermore, the prospective readers should play an important part in determining the form of presentation. Generally speaking, that method should be used with which the reader is most familiar. In this as in all other statistical work the practical aspect must not be lost sight of. If the readers are accustomed to a method and form of presentation which violates in relatively unimportant ways the principles herein suggested, that method and form should nevertheless be used; and if changes are thought to be necessary they should be instituted gradually. This is very important in practical statistical work in a business organization, since those to whom the data are to be presented generally have definite methods and forms which they want used. Any abrupt departure from these methods and forms is generally inadvisable.

Finally, the specific use to which the presented form is to be put will greatly affect the character of the presentation, particularly the size of the form, the number of copies to be made, and the nature of the impression. If the presentation is to be used in conference or in addressing a large audience, a wall chart may be used or numerous small copies may be provided for distribution. If it is to be put into a binder with other presentations or as one of a series, it should conform in size. If it is to be printed or otherwise reproduced, we usually want to avoid the use of colors, to make the impression very distinct by backing with carbon, or by using heavy black ink, to have the size properly proportioned, and so forth. Considerations of this kind are important and should always be kept in mind in the presentation of statistical data.

(For references see page 263.)

CHAPTER XXIV

THE TEXTUAL METHOD

Definition.—By textual presentation is meant the presentation of data in paragraph form. Frequently, textual discussion accompanies the statistical table or graph for the purpose of emphasizing the outstanding features of the presentation or particularly to indicate important conclusions to be drawn from the data. The distinction between this and the presentation of data in textual form is apparent. In the former case the data do not appear in the text except as their repetition enters into the discussion. In the latter, the data appear only in the text and do not appear in any other form. In the textual *discussion,* the primary aim, generally, is emphasis, clarification, or interpretation of data that appear in tabular or graphic form; in the textual *presentation* the primary aim is either to set forth the data without comment or discussion or to weave them into a descriptive or textual discussion.

Forms of Textual Presentation.—In the textual presentation the data may be, (1) woven into an interpretative or descriptive discussion, or (2) presented in solid running text. The first type may be illustrated by the following taken from page 221 of the *Economic World* for February 16, 1924:

> The steel industry is most frequently mentioned as an example of the industries highly organized and having control over prices. The price-tables of the Bureau of Labor, Washington, D. C., are based upon monthly prices of 404 commodities, divided into nine groups. For convenience in comparison the price changes are calculated in index numbers which represent percentages of the 1913 prices of the same commodities, which are reckoned as 100. The lowest average price for all these commodities in any calendar year since the war, was the year 1921, in which the index number for the average of the 404 commodities was 147. In that year the index number for the Farm Products group was 124, and for the

Metals and Metal Products group 129. The lowest point touched by the Metals and Metal Products group, however, was in 1922, when the index number was 122 and the Farm Products number was 133. The Metals and Metal Products group includes copper, lead, zinc, and silver, but in the main is made up of the principal iron and steel products. This group touched a lower average of prices in 1922 in comparison with the price level of 1913 than the Farm Products group has touched in any year since the war.

The following quotation taken from page 229 of the same publication will illustrate the second type:

The round value of the merchandise imported into the United States in January, 1924, is given by the Bureau of Foreign and Domestic Commerce as $299,000,000, compared with $288,109,624 (revised figure) for December, with $291,333,346 for November and with $329,253,664 for January, 1923. The round value of the merchandise exported in January this year was $394,000,000, against $426,798,981 (revised figure) for December, $401,480,922 for November and $335,416,506 for January, 1923. It results from these figures that the merchandise balance of trade in favor of the United States in January, 1924, was as a round amount $95,000,000, as compared with a final favorable balance of $138,689,357 in December, with a similar balance of $110,147,576 in November, but with a favorable balance of no more than $6,162,842 in January, 1923.

Data Woven into Textual Discussion.—In the form of textual presentation in which the data are woven into the textual discussion, the data do not appear in systematic order. They are introduced in the text at such points and in such ways as may be required. The different items may or may not be in juxtaposition. The primary method—if indeed it is not the only method—of emphasizing certain data, effecting comparisons, or bringing out associations and relationships, is the composition into which the figures are thrown. The general arrangement of the data and the order in which they appear depend on the nature of the discussion. The textual composition is the only instrument which is used in effecting the presentation.

Solid Running Text.—When the data are presented in solid running text the situation is quite different. A systematic order of arrangement may be observed, definite classifications maintained in proper relationship, and special emphasis given to items or groups of items as required. Thus main and sub-classes may be listed, totals and their component items properly associated, different emphases accomplished by rearranging the data in the paragraph, and so forth.

The foregoing will be clarified by the following textual presentation of the population of New York, Chicago, and Pittsburgh for the years 1900, 1910 and 1920, distributed by males, females, and total:

> The population of New York was in 1900 males, 1,705,705; females, 1,731,497; total, 3,437,202; in 1910 males, 2,382,482; females, 2,384,401; total, 4,766,883; and in 1920 males, 2,802,638; females, 2,817,410; total, 5,620,048: Chicago in 1900 males, 863,408; females, 835,167; total, 1,698,575; in 1910 males, 1,125,764; females, 1,059,519; total, 2,185,283; and in 1920 males, 1,369,917; females, 1,331,788; total, 2,701,705: Pittsburgh in 1900 males, 232,313; females, 219,199; total, 451,512; in 1910 males, 273,589; females, 260,316; total, 533,905; and in 1920, males, 296,260; females, 292,083; and total, 588,343.
>
> Or the population of New York, Chicago and Pittsburgh was for 1900, New York City, males, 1,705,705; females, 1,731,497; total, 3,437,202; Chicago, males, 863,408; females, 835,167; total, 1,698,575; and Pittsburgh, males, 232,313; females, 219,199; total, 451,512: for 1910, New York, males, 2,382,482; females, 2,384,401; total, 4,766,883; Chicago, males, 1,125,764; females, 1,059,519; total, 2,185,283; and Pittsburgh, males, 273,589; females, 260,316; toal, 533,905: for 1920, New York, males, 2,802,638; females, 2,817,410; total, 5,620,048; Chicago, males, 1,369,917; females, 1,331,788; total, 2,701,705; and Pittsburgh, males, 296,260; females, 292,083; and total, 588,343.

In the first paragraph the city as the main classification is sub-divided into years which in turn are sub-divided into males, females, and total. In the second paragraph the year as the main classification is sub-divided into cities which are in turn sub-divided into males, females, and total. Further rearrangement

might present the sex distribution as the main classification with the cities and years as sub-classes. Generally speaking, any arrangement which can be made in tabular presentation can be accomplished also in this form of the textual presentation. Likewise, special emphases can be brought out by means of distinctive type and position in this form as in the tabular presentation. Certain or all of the items in the paragraphs given above, for example, might be expressed in bold type or italics. In this form words, phrases, and punctuation marks take the place of lines in the tabular form.

Generally Ineffective.—The textual method of presentation is generally cumbersome and ineffective. This is particularly true in the running text form, especially in cases where the number of separate individual items is large, or where numerous relationships, associations, or comparisons are to be brought out. Fancy presenting the detailed population figures of New York or Chicago in textual form. It could be done and with little difference perhaps in the time and money required as compared with presenting the figures in tabular form, but for convenience the textual form would compare very unfavorably with the tabular form. In fact to cull the detailed classifications of such data from a textual presentation would be almost a hopeless task. Even in internal business statistics, to present labor turnover or wage data in textual form would be very cumbersome and would greatly detract from their serviceableness.

May Render Distinct Service.—Nevertheless, the textual method may render a distinct service in statistical presentation. The tabular and graphic methods are relatively expensive and frequently the textual method can be substituted for them with good results.

From the practical point of view, the use of the textual method should be confined for the most part to the form of weaving the data into the discussion. The presentation of data by textual method for purposes of ready reference is not to be recommended. It may be more economical in some cases and may be selected for this reason. This is rarely ever the case in practical statistical

work in a business concern and consequently the solid running form of the textual method will be of little service.

As a matter of practice, the use of the textual method in connection with the statistical service of a business organization should be confined to cases where the data are few. Its most extensive use is in the presentation of computed or derived figures such as percentages, averages, totals, and so forth, where these are but few in number. A brief memorandum, for example, may be developed to present the percentage change in sales, labor turnover, or payroll, for example, in which the figures are given at the proper points in the descriptive, explanatory or interpretive discussion.

Specifically, the most serviceable use of the textual method in practical work will be made in connection with routine tabular or graphic reports. Suppose, for example, it is desired to present periodically to the management data designed to reflect the wage level of the community for certain occupational classes comparable to those within the company. These statistics will generally be presented monthly, quarterly, or semi-annually in tabular form supplemented by graphic presentation, the data usually being in the form of averages and percentages. If, at any given date, it is observed that the changes in the data from the next preceding analysis are relatively few and unimportant, the textual method should be substituted for the tabular and graphic methods normally used and the data distributed confined to those which show a change. It is an inexcusable waste to develop the complete tabular and graphic forms for the entire body of data when for the most part they are essentially duplications of the data distributed at the last previous date.

The proper method of procedure in such case is to develop a brief textual discussion, stating the broad fact that there is practically no change in the general situation and pointing out in detail the few changes that have taken place. In this way the few data, generally percentages and averages, can be effectively presented with a substantial economy in time and money. Furthermore, it will be a great saving of time for the management, who would otherwise be compelled to wade through the full mass of facts. It is advisable, of course, to have the data in completed

form on file in the statistical organization and to state in the textual report that they are available for any who wish them.

The textual method is also frequently used by newspapers in the presentation of statistical data because of the greater convenience in the printing.

(For references see page 263.)

CHAPTER XXV

THE TABULAR METHOD

The Meaning of Tabulation.—As a general statement we may say that tabulation is an orderly association of two series of statistical classifications, one extending horizontally to the right from a given point to be read from left to right, and the other extending vertically downward from the same point, to be read from top to bottom, the proper entries in the classes of which throw the data systematically into horizontal and vertical columns. This definition is designed merely to indicate the essential elements of tabulation. The detailed characteristics which are important in practical work will be apparent in the further discussion of the subject.

It should be noted that mere listing of figures in columnar forms may not fulfill the requirements of tabulation. The tabular presentation of data may have to take into account the associations and emphases necessary to bring out their significance. This is frequently the important aspect of the tabular process. Thus two tables may be constructed from identical data to serve a particular purpose, the one being relatively worthless from the practical point of view and the other very serviceable. The one merely fulfills the requirements of the mechanics of tabulation; the other presents the data specifically to reflect their significance in the light of the purpose to be served.

This aspect of the tabular process is highly important and must be kept constantly in mind in practical tabulation. It is not enough merely to set up two series of classifications so as to produce columns of data. Very often the most important part of the tabular development is to bring out the emphases, relationships, comparisons, and associations which are significant in the given case.

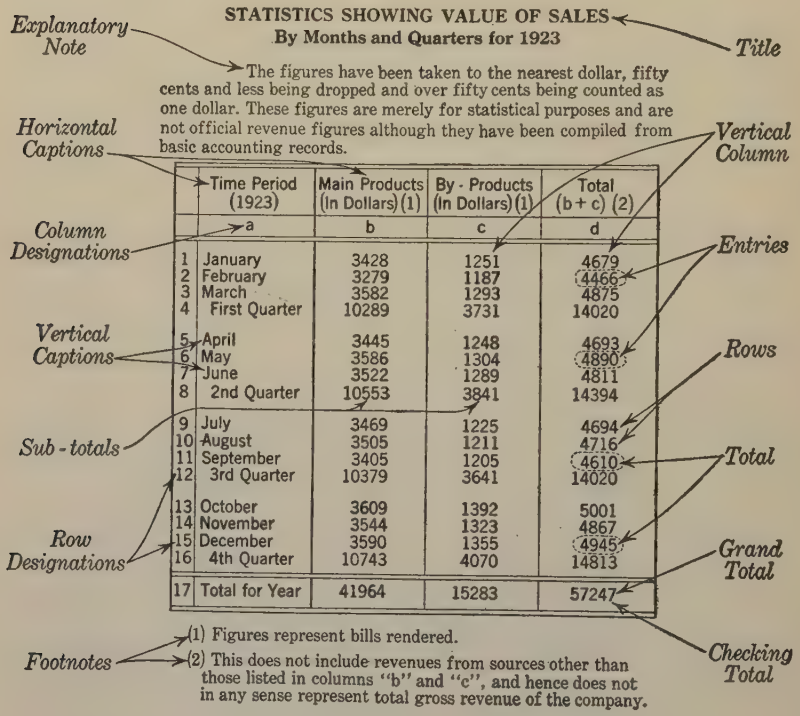
As was noted above, the tabular process is sometimes confused with analysis and interpretation although it is essentially a method

of presentation. It presupposes analysis and interpretation and takes the data which have resulted from these processes, puts them up in orderly form, and by means of certain mechanical devices emphasizes features which may be important in serving a given purpose. In the process of tabulation neither classifications nor computations are made. These are involved in working up the data. The tabular process consists merely of taking the data as they are found and presenting them in accordance with certain more or less accepted principles.

What the Table Consists of.—To make the foregoing description of the concept of tabulation more concrete, let us consider briefly the various component parts of the statistical table. The purpose here is merely to enumerate its principal parts. Their significance will be fully brought out in the next chapter. The items listed are illustrated in Table 1.

1. *Horizontal Classification*—A series of classes arranged horizontally to be read from left to right.
2. *Vertical Classification*—A series of classes arranged vertically to be read from top to bottom.
3. *Title*—This term carries the same significance here as in connection with the statistical graph (see Chapter XXIII, page 235). It may consist of the main title alone or one or more sub-titles may be required also.
4. *Explanatory Notes*—This has reference to explanations or descriptions of general features of the table which are important in the interpretation or use of the data contained therein.
5. *Footnotes*—These refer for the most part to specific points in the table as distinguished from its general aspects referred to in explanatory notes.
6. *Captions*—These are the specific designations of the several vertical and horizontal classes. Some are sub-divisions of others. They are sometimes called "Headings" and "Sub-Headings," the former referring to the main classes and the latter to their sub-divisions. Likewise the vertical classes are sometimes designated as "stubs" and the horizontal classes as "boxes."
7. *Entry*—An entry in a statistical table is a single figure which appears in the body of the table as being included in either a single one of the vertical or horizontal classes or both. The data recorded in the table are consequently referred to as entries.

8. *Column*—This refers to the list of entries included in a single final sub-caption or class of the horizontal series.
9. *Row*—This refers to the list of entries included in a single final sub-caption or class of the vertical series.
10. *Totals*—The statistical table frequently contains what are called *totals, sub-totals,* and *grand totals.* As these terms are relative it is somewhat difficult clearly to define or distinguish them. It is sufficient to say that the sub-total is a sub-division of a total which is in turn a sub-division of a grand total. They are for convenience only and generally it will not be necessary to distinguish them in practical work.



STATISTICS SHOWING VALUE OF SALES
By Months and Quarters for 1923

The figures have been taken to the nearest dollar, fifty cents and less being dropped and over fifty cents being counted as one dollar. These figures are merely for statistical purposes and are not official revenue figures although they have been compiled from basic accounting records.

	Time Period (1923)	Main Products (In Dollars) (1)	By-Products (In Dollars) (1)	Total (b + c) (2)
	a	b	c	d
1	January	3428	1251	4679
2	February	3279	1187	4466
3	March	3582	1293	4875
4	First Quarter	10289	3731	14020
5	April	3445	1248	4693
6	May	3586	1304	4890
7	June	3522	1289	4811
8	2nd Quarter	10553	3841	14394
9	July	3469	1225	4694
10	August	3505	1211	4716
11	September	3405	1205	4610
12	3rd Quarter	10379	3641	14020
13	October	3609	1392	5001
14	November	3544	1323	4867
15	December	3590	1355	4945
16	4th Quarter	10743	4070	14813
17	Total for Year	41964	15283	57247

(1) Figures represent bills rendered.
(2) This does not include revenues from sources other than those listed in columns "b" and "c", and hence does not in any sense represent total gross revenue of the company.

TABLE 1

Aim of Tabulation.—As is pointed out above, tabulation is essentially a method of presentation. Its primary function, therefore, is to set forth data in the most serviceable form for the

particular purpose in mind. The purpose may be merely to facilitate further use of the data as in the case of the data sheet in the construction of graphics or of basic records prepared in simple tabular form to make the data more accessible. It may be to facilitate analysis as is usually the case in determining classifications, computing averages, developing index numbers, and working out percentages or ratios. Or it may be to serve as a basis for interpretation as is illustrated by the completed statistical study. In any case the tabulation is purely for the sake of utility. It is used to gain an advantage in statistical work and it must take us toward the end in view. This practical aspect of the aim of tabulation must be kept in mind throughout every step of the construction of the table.

Types of Tables.—In classifying statistical tables a number of different bases have been used and a number of different results obtained. From the practical point of view classification of statistical tables on the broad basis of the purpose to be served is satisfactory. On this basis statistical tables may be classified as (1) *informative,* and (2) *analytical or interpretive.* This classification is based primarily on the broad classes of purposes to be served in tabulation as outlined in the section above. It has a really practical value. We have gained a great deal when we have classified a proposed tabular construction as *informative,* or *analytical or interpretive.* These two types differ in very essential and fundamental points of tabular construction.

It should be noted that the form of the data is not an essential element of this classification. Original detail or the most highly computed or derived figures may appear in both classes of tables.

Informative Tables.—As is pointed out above, one of the primary purposes or functions of tabulation is to set the data forth in systematic form so that they will be conveniently at hand for further use—to present them merely for information. A statistical table which is designed to serve this purpose is an *informative* table.

In this type of table no recognition is given in the tabular process to the relative significance of the data. Generally speaking,

no part of them is more important than any other part. No comparisons are to be presented, no emphases to be made, and no relationships or associations to be brought out. The sole aim is to make the data readily available in systematic and convenient form.

It is clear, therefore, that the problem of setting up a table of this type is quite different from that in connection with a table which requires emphases, comparisons, and relationships. It is

BASIC SALES DATA
November, 1924

Sales Order	Date of Sale [Day of Month]	Place	Sales-man	Class of Prod-uct	Grade of Prod-uct	Size of Prod-uct	Quan-tity of Goods	Value of Goods	Price	Name of Customer	Class of Customer (1)
a	b	c	d	e	f	g	h	i	j	k	l
1											
2											
3											
4											
5											
6											
7											
8											
9											
10											
11											
12											
13											
14											
15											
16											
17											
18											
19											
20											

[1] Dealer, Jobber, Private.

TABLE 2.—A Typical Informative Type of Table

generally simpler in form and its construction is much less difficult. The more difficult problems of tabulation are not involved in this type of table. Table 2 will serve as an illustration. The distinction between the two classes of tables will be further clarified in the discussion of analytical or interpretative tables.

Analytical or Interpretative Tables.—The analytical or interpretative table is designed primarily to bring out the significance

of the data presented. The significance emphasized may or may not be that of the data in their relation to a specific problem. It may have reference merely to the general analysis of the data. The purpose may be to bring out important features of the data which have been revealed through analysis or it may be merely for the purpose of revealing significant aspects of the data for further analysis.

It is evident, therefore, that this type of table presents a much more complicated problem than the informative type. It differs from the informative primarily on two points. In the first place, it aims to place definite emphasis on certain entries or groups of entries and the various devices in tabular construction are brought into play to accomplish this end. In the second place, various comparisons, relationships, or associations are to be portrayed by means of the tabulation. These two points of distinction are very fundamental and important. Their accomplishment involves the most difficult problems of tabulation. All of the "tricks of the trade" are called into play in the construction of this type of table. Table 1 on page 248 partakes of the nature of an analytical table.

Construction of the Statistical Table.—Thus far in the discussion of tabulation only the general aspects of the subject have been considered. What has been said will be only generally helpful in the actual work of tabular construction. It is, however, a sort of background or foundation on which the detailed features of the process of tabulation may be developed. The next chapter therefore, will be devoted to a discussion of the detailed methods of procedure in working up a statistical table.

(For references see page 263.)

CHAPTER XXVI

CONSTRUCTING A STATISTICAL TABLE

Steps in the Process.—It is difficult to outline in sufficiently concrete form to be of practical service the successive steps to be followed in the construction of a statistical table. The development of the several parts of the table as outlined in Chapter XXV, page 247, is carried on to a greater or less extent concurrently. No one part is completed independently before passing to another. Generally, the development of one part contributes substantially to the completion of other parts.

There are, however, certain broad steps which, in the main, are followed in building up a statistical table. These are as follows:

1. Define the purpose.
2. Select the main vertical and horizontal classes and sub-classes and form the headings and sub-headings.
3. Draw rough draft of proposed forms, thoroughly inspect them, and thus select the proper form.
4. Select methods of emphasis to be used.
5. Enter data, reinspect the form, and put body of table in final shape.
6. Select and enter the title.
7. Select and enter general explanatory and qualifying notes and footnotes.
8. Make final inspection for accuracy and completeness.

These various steps will be discussed in order.

Defining the Purpose.—It will be recalled from previous discussions that the purpose greatly affects if it does not control the tabulation. The same data will be put up in one tabular form to serve one purpose and in another form to serve another purpose. The arrangement of the classes of data, the general form of the table, the devices used for emphasis, and even the title and explana-

tory notes will be determined only after they have been fully and thoroughly considered in the light of the purpose.

Defining the purpose at the outset will also enable us to determine the number of tables required and the size that the finished table should take.

Selecting the Main and Sub-classes.—It will be recalled that the classification of the data will have been completed before tabulation begins. Since the importance of the classifications will vary with the purpose to be served, it is a function of tabulation to arrange them in such a way as to reflect their relative importance. This will involve specifically the selection of the main vertical and horizontal classes and the sub-classes of each to be set up in the table. In doing this it will be well to construct skeleton outlines of tabular forms showing different class arrangements, in order to facilitate inspection and evaluation. In making these selections it should be kept in mind that generally a few main classes with proper sub-classes are preferable to several coordinate classes. In the informative type of table, however, this may not be the case: convenience in reference may be better served frequently by listing the classes coordinately.

The problem of selecting the classes is a two-fold one, involving primarily (1) the determination of which series of classes will be placed vertically and which horizontally, and (2) the arrangement of the classes within the series.

In regard to the first of these, definite directions cannot be given which will apply comprehensively to all forms of tables. In the tabulation of variable data, the independent variable is placed vertically and the dependent variable is placed horizontally. In tabulating historical data, for example, the time periods are usually given the vertical position as is illustrated in Table 1, page 248, and in frequency data, the measurements are usually given the vertical position and the frequencies the horizontal position as shown in Table 1 of Chapter XXXV. However, there will be variations from this practice to suit the requirements of the given case, particularly in tabulating historical data.

It is sometimes asserted also that the vertical position is more

important. The statement is based for the most part on the fact that figures of a series arranged in vertical column are more readily inspected and compared than those arranged in horizontal column. Practically, however, the difference in the importance of the two positions is not sufficient to be a determining factor in every case. Convenience of space, for example, frequently outweighs the factor of relative importance.

The Order of Arrangement.—In selecting the main and sub-classes careful attention should be given to the arrangement of the classes. This is a very important aspect of the tabular process. It will apply to both series of classes although it is most effective in connection with the vertical.

KILOWATT HOURS OF ELECTRICITY CONSUMED
Year 1924

	Geographical Section	Absolute Amount (000 omitted)	Per Cent of Total
	a	b	c
1	West South Central	708,400	2.96
2	East South Central	830,000	3.47
3	West North Central	1,371,000	5.74
4	New England	1,424,000	5.96
5	Mountain	1,500,000	6.28
6	Pacific	3,251,000	13.61
7	South Atlantic	3,383,000	14.16
8	Middle Atlantic	5,704,000	23.87
9	East North Central	5,722,000	23.95
10	Total	23,893,400	100.00

TABLE 3.—Size of Items the Basis of Arrangement

Note the use of "000 omitted" to indicate that the figures in column b are in thousands. The term "in thousands" might be used instead. Note also that the rows are designated by numbers placed just to the left of the vertical classes and the columns by letters placed just below the horizontal classes.

The significance of this aspect of tabulation will be apparent from the discussion of the following methods of arrangement:

1. *Alphabetical*—In this case the classes are arranged in simple alphabetical order. This arrangement is particularly serviceable in reference tables.
2. *Order of Size*—The numerical value of the items may be the determining factor in the arrangement. It will apply generally to the frequency classifications where the quantitative groups are arrayed as shown in Table 3, page 384. Or the classes may be arranged so that the entries may be in ascending or descending order as illustrated in Table 1. Generally speaking it will be practicable only in connection with a table having a single vertical column of entries or additional columns derived from it. It will be serviceable particularly where comparisons are to be made.
3. *Chronological*—Again chronology may be the controlling factor and the classes arranged in order of time. Table 1, page 248, illustrates this type of arrangement. It will be used generally in any presentation of historical data. It should be noted in this connection that the plan of reversing the order by placing the later time periods at the top and left of the table is not to be recommended as a general practice.
4. *Relative Position*—In other cases, the relative positions of the classes may determine the arrangement. In presenting data by certain territorial units such as states, counties, or sales districts, the territorial units may be arranged according to their nearness to one another, a certain one having been selected as a starting point. Table 4 is typical of this order of arrangement.

   Likewise, in tabulating internal data of the company by departments, the departments may be arranged in order from the executive department down or vice versa: or they may be listed in order of the processes of the company, such as purchasing, producing and marketing.
5. *Causal Relationship*—Finally causal relationship between the several classes may be the basis of the arrangement. In the analytical or interpretative type of table this will be frequently the most effective arrangement. The classes will be grouped to bring out the required comparisons, relationships or asso-

ciations even though none of the arrangements suggested above may serve. In this connection it should be emphasized, that both columns and entries to be compared, associated, or related should be placed as near each other as is feasible.

In any event, the purpose to be served will be the fundamentally controlling factor. Each of the different methods of arrangement will be evaluated in the light of the purpose of the

KILOWATT HOURS OF ELECTRICITY CONSUMED
Year 1924

	Geographical Section	Absolute Amount (000 omitted)	Per Cent of Total
	a	b	c
1	New England	1,424,900	5.96
2	Middle Atlantic	5,704,000	23.87
3	South Atlantic	3,383,000	14.16
4	East North Central	5,722,000	23.95
5	West North Central	1,371,000	5.74
6	East South Central	830,000	3.47
7	West South Central	708,400	2.96
8	Mountain	1,500,000	6.28
9	Pacific	3,251,000	13.61
10	Total	23,894,300	100.00

TABLE 4.—Relative Position the Basis of Arrangement

tabulation and that one chosen which is most effective in accomplishing it. Our aim here is merely to emphasize the importance of thoroughly considering the order of arrangement of the classes and to suggest the arrangement most frequently used. Whatever the order of arrangement may be, the columns or entries that are to be compared or related in any special way should be associated as closely as possible.

Rough Working Draft of Table.—When the relative importance of the several classes of data has been determined as suggested in the preceding section, a rough draft of the table should be formed. This should show the main horizontal and vertical classes and the subdivisions under each. Roughly phrased captions should be entered, thus making the body of the table complete except for the entries.

This rough draft will be serviceable in different ways. It aids in visualizing the finished table; it serves, when completed, to send to the typist; and, it facilitates the selection of the specific form of table. Furthermore, on the basis of the rough draft the width of the columns can be determined and the total amount of space vertically and horizontally that will be required. In this connection it may be noted that it is generally advisable to standardize the size of statistical tables in the company so far as possible, 8½ by 11 being a convenient size. If this practice is followed, however, a great deal of attention will be required frequently to get the tabular form to fit into the size and shape of paper. In no case should the effectiveness of the presentation be sacrificed to mechanical convenience.

It will be observed also that the vertical classification should be placed at the left of the tabular form to be read from top to bottom and that the horizontal classification should be placed at the top of the tabular form to be read from left to right. In other words, the top of the vertical classification and the left end of the horizontal classification should join at the upper left corner of the tabular form. The practice observed in reading and writing textual matter is followed here—the procedure is from top to bottom and left to right. The vertical classification may be repeated at the right if required in any given case. It is not often necessary to repeat the horizontal classification.

On the basis of this rough draft the main and sub-classes should be thoroughly reconsidered and any adjustments or improvements made which may seem advisable.

Methods of Emphasis in Table Construction.—In addition to the general character of the tables it is frequently important to emphasize certain individual entries or groups of entries in

order better to serve the purpose in mind. This is particularly true in the analytical or interpretative type of table. Generally in analyzing data with respect to their significance for a specific purpose, certain of them will be of greater importance than others. This relative importance as well as any comparisons, associations, or relationships must be brought out in tabulation.

This may be done in a number of different ways, chief of which are the following:

1. *Distinctive Type*—This is a favorite method of emphasizing outstanding entries. The type may be distinctive in color, size, or form. Red figures are freely used to show decreases. Sometimes heavy type as offsetting light type or italics as offsetting plain type are used to make entries stand out. Distinctive type is used particularly to emphasize outstanding percentages, averages, ratios and so forth which depart from the average, normal or standard.
2. *Distinctive Lines*—Likewise the use of single and double, light and heavy, or colored lines serve the same purpose. This is particularly true where groups of items are to be emphasized. Important percentages, averages and totals, may be set off by double, heavy, or colored lines. Likewise a caption of particular importance may be so distinguished.
3. *Position*—The importance of position may be considered from two angles. In the first place the classifications or captions have varying importance. The vertical and horizontal classifications are sometimes referred to as primary and secondary respectively, the former being considered as the more important position. It is contended, therefore, that the vertical classification, generally speaking, tends to give a comparative emphasis. In practical work, however, it is allowable to shift freely from one to the other for convenience in the arrangement of the data.

   Furthermore, different positions within the table carry different degrees of emphasis. It is generally assumed that the positions nearest the captions are more important and the upper left-hand corner consequently the most important of all. Items placed in this corner, it is assumed, are emphasized because of their nearness to the two sets of captions. Moreover, within a certain horizontal or vertical column the

positions next to the captions are the most important and importance gradually diminishes as the distance therefrom increases. It may be remarked, however, that it is easy to spend too much time and attention on the relative position of items in a table. The relative importance of the several positions is generally over-emphasized. As a matter of practice, very little is gained by fine distinctions of this character.

4. *Order of Arrangement*—The order of arrangement of the items in the vertical and horizontal classifications may also be used as emphasis. As stated above, various factors may serve as the primary basis of this arrangement. If chronology is the controlling factor, the time element is emphasized; if size, the relative value of the items is emphasized; if causal relationships, this factor is emphasized.
5. *Special Markings*—Individual entries may be emphasized through special markings such as a circle around the entry or a line underscoring it.

All of these things should be freely resorted to in accomplishing the emphasis required in the particular case. In addition, the title, explanatory notes, footnotes, captions and sub-captions, will also be serviceable in bringing out special emphasis. These are more serviceable, however, in defining the general character of the table and are discussed fully on pages 235 to 238.

Entering Data.—In most cases, the data should be entered in the rough draft of the table and thoroughly inspected. Great care should be taken to see that these entries are properly made and that the data are accurate in every respect.

It is important to note, however, that frequently if not generally, in practical work in a business concern the data will be transcribed directly from the work sheets to the tabular form by the typist and will not be entered at all in the rough draft of the table. Whenever the time will permit, it is much preferable actually to enter them in the table. This will afford an opportunity to inspect the data carefully before they are typed finally. It will also reduce the typing to straight copy work and thus lessen the chance of error in this process.

Putting Table in Final Shape.—Before the rough draft is submitted for final typing or printing it should be put in final shape. This will involve the addition of certain minor mechanical features which, though not vital to the tabulation, nevertheless contribute substantially to its effectiveness. The following are typical of these features and are offered primarily as suggestions:

1. *Rulings*—The main parts or sections of the table, if any, should be separated by double, heavy, or colored lines. Major totals or other principal items or groups of items may also be so distinguished. Likewise, distinctive lines should be placed at the extreme top and bottom of the table. The double and heavy black lines are more practicable than colored lines particularly if the table is to be reproduced.

   Light lines drawn across the table horizontally at regular intervals will lend effectiveness to the presentation by aiding the eye to follow across the page. The same may be accomplished by grouping entries in the vertical columns into groups of five, eight, ten, or other number. These devices are illustrated in Tables 2 and 1, pages 250 and 248, respectively.
2. *Numbering the Columns*—Both the columns and rows should be numbered. The columns should be numbered at the top just below the captions—*i.e.*, the column numbers should be placed between the final sub-captions and the entries in the body of the table and should be set off with light lines. This position is necessary in order that the numbers may be conveniently associated with the final sub-captions. If it is desired in any given case to designate the main classes in the horizontal series, the numbers or letters for the purpose may be placed above the captions and likewise set off with light lines and for intermediate classes they may be placed in the box with the heading.

   The rows will generally be numbered at the left of the table. These numbers may be placed to the right or left of the captions, depending on the character of the captions. If the captions and sub-captions are associated in the way the column captions are usually associated, the numbers will be placed just to the left of the final sub-captions and set off by light lines, as shown in Table 5. Usually, however, the captions and sub-captions in these classifications are arranged as shown in Figure 134 Appendix B, and in such

case the numbers will be placed to the left of the captions as indicated in the same table.

If the table is large, the rows should be numbered at both sides of the table. Rarely will it be necessary to repeat the the numbers of the columns at the bottom of the table.

Letters may be used for this purpose instead of numbers. A convenient scheme is to use letters for the columns and numbers for the rows. This scheme is used in Table 4 above.

STATISTICS OF ANTHRACITE FOR HOUSEHOLD USE
AVERAGE RETAIL PRICE PER TON OF 2,000 POUNDS
AT BOSTON, NEW YORK, AND CHICAGO
1913, 1923 and 1924

Cities		Grades		1913		1923	1924	
				Jan. 15	July 15	Oct. 15	Sept. 15	Oct. 15
a		b		c	d	e	f	g
I	Boston	1	Stove	8.25	7.50	16.00	16.00	16.00
		2	Chestnut	8.25	7.75	16.00	16.00	16.00
II	New York	3	Stove	7.07	6.66	14.58	14.13	14.21
		4	Chestnut	7.14	6.80	14.58	14.13	14.17
III	Chicago	5	Stove	8.00	7.80	17.00	16.50	16.75
		6	Chestnut	8.25	8.05	17.00	16.50	16.75

TABLE 5.—Illustrating Method of Numbering the Horizontal Columns or Rows

These designations are particularly serviceable for reference in textual discussion, explanatory notes, and so forth. They are also convenient in making definite the caption headings as illustrated in Table 1, page 248: column "d" shown to be column "b" plus column "c". Furthermore, if more than one page is used in the table, it may be satisfactory merely to repeat the numbers of the column headings and not repeat the wording. This will of course be somewhat of

a hindrance in reading the table, but the difficulty may be offset by the resultant economy of time and space in the construction.

Title; Explanatory Notes; Footnotes.—These steps in presentation are discussed in detail on pages 235 to 238. Their position is the only point that need concern us here.

Generally, the preferable position for the title is at the extreme top. The explanatory notes are likewise preferably placed at the top, between the title and the body of the table. Both of these, however, may be located at the bottom of the table, the title being placed next to the body of the table and the explanatory notes below it. As a matter of practice, lack of space may require that the explanatory notes be placed on the reverse of the tabular form or in a separate memorandum.

The footnotes, which are applicable to special points of the table, should be associated as far as possible with the particular points to which they apply. In practice, the preferable method is to place the footnote below the table and carefully to indicate in the text each item to which it refers.

The fundamental principle here is that the practices observed in reading textual matter should be followed in the statistical table —*i.e.*, in reading the table we should proceed from the top downward and from left to right.

Final Inspection for Accuracy and Completeness.—When the table has been finally completed, it should be carefully inspected and checked in its every part. Beginning with the title, each of the component parts listed on page 247 will be taken step by step and given the most careful scrutiny. This should be done by some one who is thoroughly familiar with the data and who appreciates their significance. Reference is made here to checking over and above that which is regularly done during the compilation and entering of data.

The methods of checking which are briefly outlined in Appendix E, Part 5, will be serviceable here. Some will be especially serviceable in checking the title and captions, others in checking the computations and entries, and still others in checking the mechanics of the table. It is advisable to establish a standard

system of checks which are applied regularly to the statistical table.

The important point is that the table should be thoroughly checked: the detailed method of accomplishing this will be determined in each specific case. The extent to which checking can be done will frequently be limited by the time available. It is not unusual in practical work to be so rushed in the preparation of the table as not to have time for thorough final inspection.

The Practical Aspect.—In the use of the foregoing suggestions in table construction, the practical aspects must be kept in mind. From the practical point of view it is very easy to overdo the fine points of tabulation. The suggestions and directions given herein are designed to aid generally in the development of the ideal tabular presentation. They are, however, mere suggestions and if the amateur finds in actual practice that there are violations of these suggestions, he will certainly be prudent enough not to condemn the tabular presentation because of these minor points. If he considers them important, he can improve the work by gradual effort; but by all means, if the essential elements of the purpose are being served, then no interference should be made merely to correct certain detailed features whose importance are more academic than practical.

The table, for example, should have distinctive lines at the top and bottom, but it would be foolish to say that a table is useless if it should not happen to have distinctive lines at the top and bottom. The same may be said for a number of other suggestions that are offered. A final test of a table is its contribution to the purpose in view and frequently this may be well served without fine observance of the more detailed suggestions given in these discussions.

References—Chapters XXIII to XXVI (Textual and Tabular Presentation)

Bowley, A. L. Elements of Statistics, 2nd Ed., 1902, Ch. IV.
Bowley, A. L. An Elementary Manual of Statistics, 1910, Ch. VI.
Davies, G. R. Introduction to Economic Statistics, 1912, Ch. I.
Jerome, Harry. Statistical Method, 1924, pp. 28-49.
King, W. I. Elements of Statistical Method, 1912, Ch. IX.
Secrist, Horace. An Introduction to Statistical Methods, 1917. Ch.V.

CHAPTER XXVII

THE GRAPHIC METHOD—GENERAL ASPECTS

The Meaning of Graphics.—Some effort has been made in statistical literature to confine the meaning of the term "graphics" to a certain limited conception, distinguishing it from other closely related terms. Whatever may be said for such efforts from the theoretical or academic point of view, the persistent use of the term "graphics" in the field of business as covering every device of this kind indicates that the proper conception of the term from the practical point of view is to make it fully comprehensive and inclusive of all types and forms of graphic representation. From this point of view the concept may be concretely described by saying simply that graphics is the *method of presentation in which the statistical data are represented by lines or bars, curves, areas, volumes, pictures or other kinds of diagrams.*

It should be noted in this connection that this concept excludes such graphic representation as organization charts and procedure or flow charts. Since it is clear that such charts are not statistical graphs in any sense of the term—since they do not represent numerical data—the true character of graphic presentation of statistical data should not be confused by a discussion of them. They belong to the more special subject of office organization and management and not to statistics.

In the present discussion this broad conception of graphics will be followed. On the other hand, the distinction will be maintained between the terms "graph" and "chart" which are generally used interchangeably. For purposes of convenience the word "graph" will be used to designate the figure plotted on the grid or chart field to represent the data—*i.e.*, the curve, bar, circle, or other form—and the word "chart" to designate the completed presentation including the title, explanatory notes, footnotes, key, and so forth.

Component Parts of the Graphic Chart.—The statistical graphs which will be used most extensively in practical business statistics are the curve and the bar graphs. Since the former is more extensively used the principal component parts of the curve chart may be enumerated at this point. This will tend to clarify the discussion of the meaning of graphics just given and will familiarize the reader with terminology required in the further discussion of the subject. From this list the component parts of other forms of graphic charts will be readily determined by inspection.

The list given below shows the principal parts of which the completed curve chart of the arithmetic type is composed. They are illustrated in Figure 23. Detailed discussion of them will be found in the next chapter.

1. *Title*—The title is quite similar to that of the statistical table. It will be composed generally of a main *title* and one or more sub-titles.
2. *The Chart-Field or the Grid*—This is the grating-like base on which the graph is constructed. The scales are graduated on the chart field and the graph is confined within its limits.
3. *Origin or Zero Point*—This is the starting point for all measurements on the grid. The measurements may extend in any direction from this point. In the most common forms they extend horizontally to the right and vertically upward.
4. *Axes*—These are the two lines, one horizontal and the other vertical, which intersect at the origin or zero point. The horizontal line is called the X-axis and the vertical line the Y-axis.
6. *Base Line*—This is a more or less technical term generally applied to those portions of the X-axis and Y-axis which the grid comprehends. The portion of the Y-axis is called the vertical base line and the portion of the X-axis the horizontal base line. In cases where the zero line or technical base lines are omitted from the grid, the coordinate lines which lie at the bottom and left of the grid are somewhat loosely called the vertical and horizontal base lines respectively.
7. *Scale*—This is a term used to designate the graduations or

calibrations on the base lines of the chart. Those on the horizontal base line are called the *horizontal scale* and those on the vertical base line the *vertical scale*. In other words, as Mr. Karsten points out on page 167 of his "Charts and Graphs," they are the figures which tell us the values that are arbitrarily assigned to linear distances, horizontal and vertical.

8. *Scale Caption*—This is the descriptive wording placed alongside each scale to designate the value of a single unit of the scale. Thus a scale for wages may represent dollars per month or year, or cents per hour.
9. *Plotted Points*—These are the points on the chart field which are determined by the intersections of the coordinate lines that represent the data on the two scales.
10. *The Graph*—This is the specific diagram plotted on the chart field to represent the data. It is the result of plotting the data on the grid or chart field in accordance with the measurements prescribed by the vertical and horizontal scales. It is the curve or other diagram which actually registers the data.
11. *Key or Legend*—When more than one graph is presented on a single chart field, it is necessary to indicate clearly the particular data which each one represents. This is sometimes done by constructing in an open space on the field, or other convenient place, short specimens of the graphs and indicating the character of data which they each represent. These specimens taken together are called a key or legend.

The Aim of Graphics.—Fundamentally, the aim of graphic presentation is to clarify or emphasize the essential facts or conclusions of statistical data. The graphic method is a device designed to bring out more effectively the story which the data have to tell. Its function is aptly portrayed in a series of articles on the subject by M. C. Rorty which appeared in *Industrial Management* for December 1920 and January and February 1921 under the very suggestive title, "Making Statistics Talk."

This fundamental aspect of graphic presentation should be kept constantly in mind in the use of graphic devices. There is some tendency today to make the graphic chart the end in itself rather than a means to an end. A sort of "graphitis" is developing

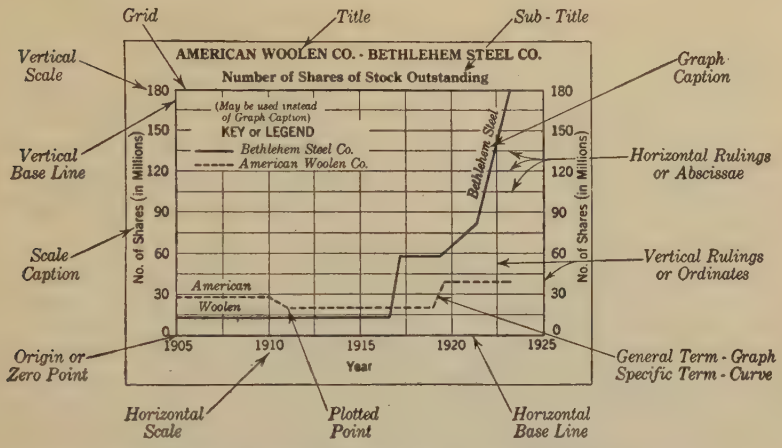
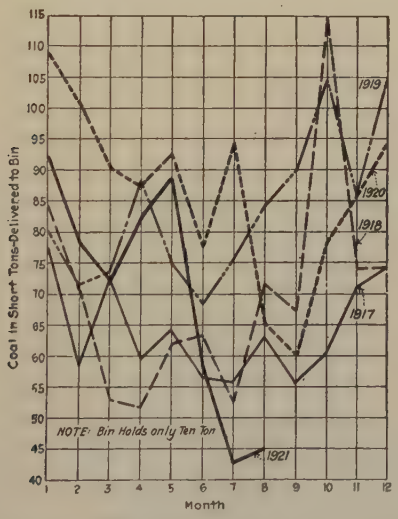


FIGURE 23.—Principal Parts of which the Graphic Chart is Composed and the Terms generally used to Designate Them

The entire figure, including all the parts indicated, constitutes a *graphic chart.*

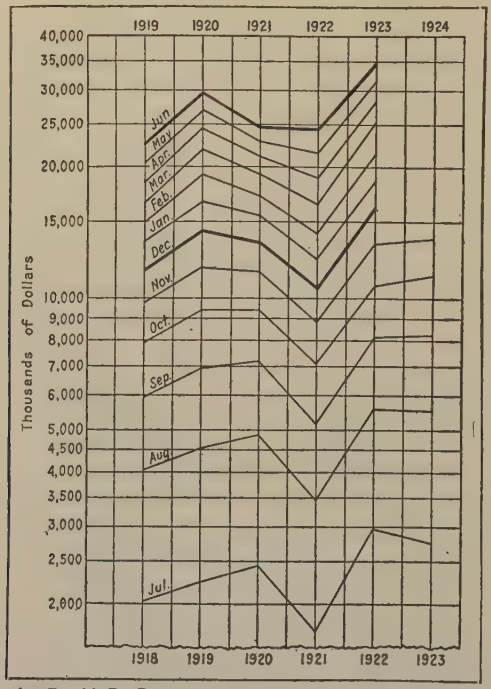


(From article by M. C. W. Tomlinson, *Management Engineering,* Dec. 1921)

FIGURE 24.—Yearly Coal Consumption Curves for Power and Heating Requirements of a Tannery.

A vague graphic presentation.

which leads to the indiscriminate construction of graphic devices without proper consideration as to the service they may render. Frequently, graphic charts are so complicated that it is more difficult to read them than to get the facts from the data in tabular form. The value of the graph in Figure 24, for example, is questionable. The aim of graphics is to *clarify* or *emphasize* and

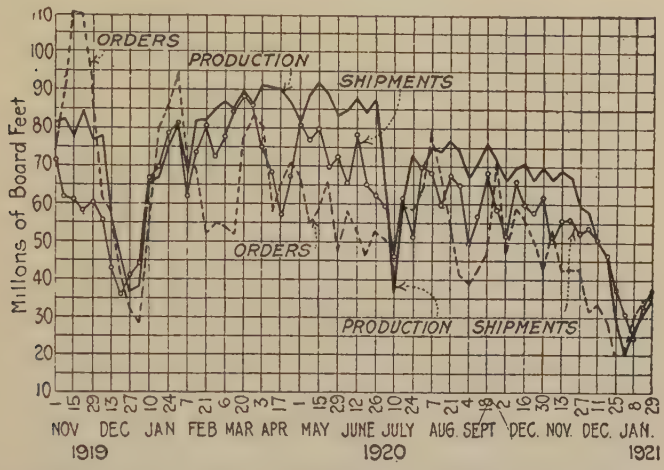


(From article by David B. Porter, *Management and Administration*, Jan. 1924)
FIGURE 25.—Cumulative Chart of Sales Allowing for Indefinite Extension. A clear presentation of many curves.

the graphic method should be resorted to only for the definite accomplishment of this aim.

In this connection, it should be emphasized that we cannot definitely indicate the maximum number of curves that should be placed on a single grid. The folly of the arbitrary statements frequently made that in no case should more than 3 or 4 curves be presented on a single grid is apparent from Figures 25 and 26.

The graphic presentation of Figure 25 with 11 curves is more effective than that of Figure 26 with only 3 curves.



(From article by P. M. Marshall, *Management Engineering*, Oct. 1922)

FIGURE 26.—Typical Chart of Production, Shipments and Orders Over a Period of Fifteen Months.

A vague presentation of but three curves.

Furthermore, in practical work data are frequently thrown into graphic form when its use is not required. This point is emphasized in the following quotation, taken from an article in the January, 1924, issue of *Management and Administration*, page 20, by H. D. Gidney, Statistician of the Pacific Telephone and Telegraph Company:

> "In closing, a warning should be sounded in regard to the furnishing of graphic charts and reports. A person working with graphics will often become so enthusiastic that he will turn out charts in the same manner as the presses turn out German marks. And the majority of the charts will be of about as much value to the executive as the mark. Care should be taken that the basic data is correct and that the chart will be of real use to the executive, and also that the cost of producing the chart is not more than the value to the executive. Attention to this warning will result in weeding out a great many of what appears at first thought to be very valuable charts."

Types or Classes of Statistical Graphs.—From the practical point of view in business statistics a simple classification of statis-

tical graphs based on the distinctive nature or appearance is quite adequate and more serviceable than the more refined or involved groupings. On this broad and more or less indefinite basis the following primary classes of statistical graphs may be established:

1. Columns and bars
2. Curves
3. Maps
4. Circles
5. Areas
6. Volumes
7. Pictures
8. Combination Graphs

CHAPTER XXVIII

FUNDAMENTALS OF GRAPHIC CONSTRUCTION

The Fundamental Fact of Graphics.—Graphic representation is based fundamentally on the simple principle of representing numerical facts by linear distances or functions thereof. In its essential aspects this principle is generally familiar. It is not fundamentally different from the scheme of the ordinary foot-rule, yard stick, tape measure and other linear measuring instruments. We are all familiar with the method of representing figures by distances on these instruments. We start at a certain point marked "o" and at a certain distance from this point we make a mark and call it 1 inch, at other distances other marks which we call 2 inches, a foot, yard, rod, and so forth. It is important to note that the graduations on these rules have been established by custom. We merely agree that a certain distance should represent a certain unit number and from this basic representation we build up the distances corresponding to other numbers which are functions of the unit. Series of graduations of this kind in graphics are known as scales.

Now, the distances or spaces graduated on these scales may represent statistical data as well as units of linear measurement. The inch space on the foot-rule shown in A of Figure 27 may just as well represent $100 of revenue, 10 employes, 50 sales or what not, as is shown in B, C, and D of the same figure. The unit space may represent any kind or form of data. It may represent whatever we may decide to have it represent. This is purely an arbitrary matter, a "rule of the game" that we set up to guide the development and use of the device. Likewise, the graduations on the line may be varied to suit the specific case as shown in Figures 28, in which the unit distances differ but the number represented remains constant.

This typifies the fundamental fact of graphic representation. We arbitrarily select a certain linear distance or area or volume to represent a unit of data and from this basic fact we build up

the graph which completely represents the data that are to be portrayed. Further discussion of the use of linear distances to represent statistical data will bring out the important phases of this fundamental principle.

The Single Scale.—On the basis of the foregoing we can very readily represent statistical data by linear distances. There

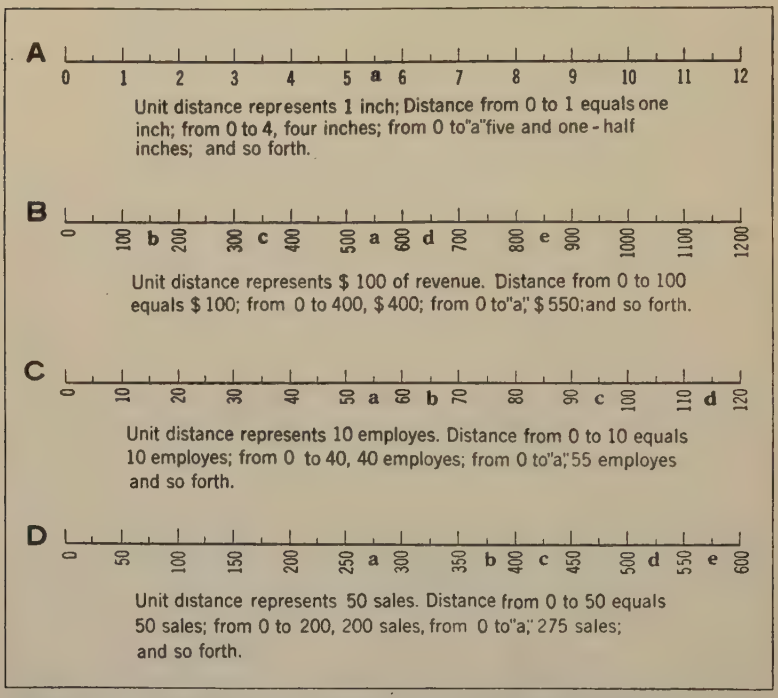

A
0 1 2 3 4 5 a 6 7 8 9 10 11 12
Unit distance represents 1 inch; Distance from 0 to 1 equals one inch; from 0 to 4, four inches; from 0 to "a" five and one-half inches; and so forth.
B
0 100 b 200 300 c 400 500 a 600 d 700 800 e 900 1000 1100 1200
Unit distance represents $ 100 of revenue. Distance from 0 to 100 equals $ 100; from 0 to 400, $ 400; from 0 to "a," $ 550; and so forth.
C
0 10 20 30 40 50 a 60 b 70 80 90 c 100 110 d 120
Unit distance represents 10 employes. Distance from 0 to 10 equals 10 employes; from 0 to 40, 40 employes; from 0 to "a," 55 employes and so forth.
D
0 50 100 150 200 250 a 300 350 b 400 c 450 500 d 550 e 600
Unit distance represents 50 sales. Distance from 0 to 50 equals 50 sales; from 0 to 200, 200 sales, from 0 to "a," 275 sales; and so forth.

FIGURE 27.—Graduations of Ordinary Foot-Rule Used to Represent Statistical Data.

The student should practice scale reading by determining the values of the points b, c, d, e and others.

are certain arbitrary rules in connection with the use of this device which are important to bear in mind:

1. A starting point is selected and marked "0," just as in the case of instruments for linear measurements. From this starting point a line is drawn horizontally to the right and graduations laid off on it. (See #6 below.)

2. The graduations on this line are not fixed as they are on linear measuring instruments. They may be varied to suit the separate requirements of different graphic problems. A given distance may represent 50 dollars of revenue, for example, in one graph, $100 of revenue in another graph, $1,000 in a third, and so forth, as is illustrated in Figure 28. The unit distance which is to represent a unit of data will be specially selected in connection with each separate graph.
3. Any kind or form of data may be represented by the same series of graduations. A given distance may represent a unit of data which reflect numbers of employes, prices of com-

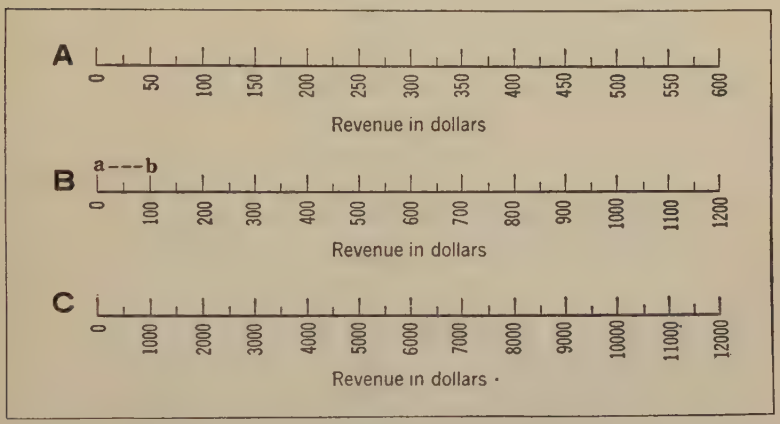


FIGURE 28.—Equal Distances on the Three Scales Represent Respectively $50 on A, $100 on B, and $1,000 on C

modities, sales of the company's goods, manufacturing output, security prices, and so forth. Furthermore, these facts may be in the form of absolute numbers or computed or derived figures, such as averages, percentages, or index numbers.

4. It is apparent that after having selected a distance to represent a unit of data we can from this fact determine the proper distances to represent data of different magnitudes. Since the distance "a" to "b" in B of Figure 28 represents $100 revenue, it is clear that the distance corresponding to any amount of revenue can be determined. $200 will be represented by twice this distance; $1,000, ten times; $50, one-half; $25, one-quarter; and so forth. The scale can be made indefinitely long, thereby making possible the representation

of numbers of indefinite magnitude. We decide that a certain distance will represent a given unit of data and then any given value will be represented by as many times the unit distance as the unit value it represents is contained times in the value to be represented.

5. *Vertical Scale*—The foregoing will be equally applicable to a vertical scale as illustrated in Figure 29.
6. It should be noted that the direction which the scale takes from the starting point is not fixed. The horizontal directions of the scales in the illustrations noted above are purely ar-

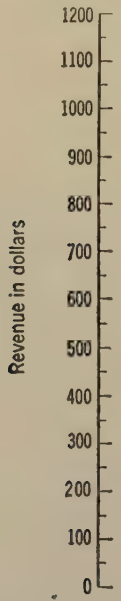


FIGURE 29.—The Scale may be Placed Vertically as well as Horizontally

bitrary. The scale may take a direction horizontally to the right or left or vertically upward or downward, as is illustrated in Figure 30, or in fact any other direction, although most practical work is confined to the vertical and horizontal scales. This does not affect the essential character of the scale. The graduations will have the same value in number of scale units in whatever direction the scale may be projected. A given distance on any of the scales in Figure 30, for example, will represent an equal number of scale units.

7. *Positive and Negative Values*—It is by means of the directions of the scales that we distinguish positive and negative values. By established practice the graduations on the horizontal scale to the left of the starting point are negative and those to the right are positive. On the vertical scale, those above the starting point are positive and those below it are negative. This practice is illustrated in Figures 30 and 31.

We are now able to plot statistical data on a simple horizontal or vertical line. After the unit distance has been selected to represent a unit of data, any number greater than zero may be plotted on the horizontal scale to the right of the starting point or on the vertical scale above the starting point and any number less

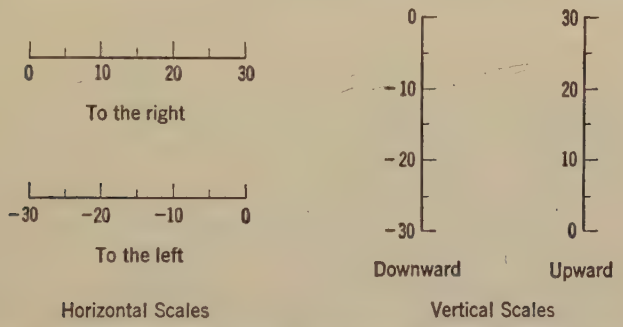


FIGURE 30.—Scales may Extend Vertically Upward or Downward or Horizontally to the Left or Right from the Zero Point or Origin

than zero may be plotted to the left of the starting point on the horizontal scale and below the starting point on the vertical scale.

The single scale is all that is required in the simple bar-graph. In this type of graph, the data are represented by a single dimension, *i.e.*, the length of the bar. So far as this type is concerned we could stop right here. In other types of graphs however, we require the use of two scales in association.

Two Scales in Association.—The device illustrated in Figure 31 is also required in graphic representation. It will be observed that this device consists merely of a vertical and horizontal scale intersecting each other perpendicularly at the starting or zero point of each. These two perpendicular lines divide the field into four

equal parts, known as quadrants. The quadrants are designated by numbers beginning with number "1" for the upper right-hand quadrant and proceeding counter clock-wise as shown in the figure.

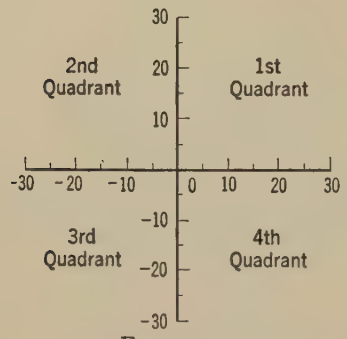


FIGURE 31

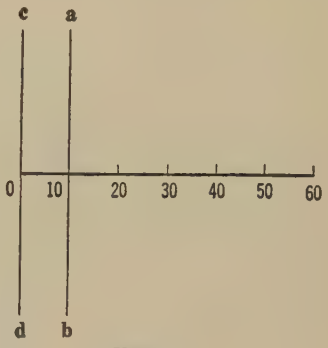


FIGURE 32

A Further Property of the Graduations on the Simple Scale.—Assume that perpendiculars are drawn through the starting point and any other point on a horizontal scale as shown in Figure 32. It is evident now that the perpendicular distances be-

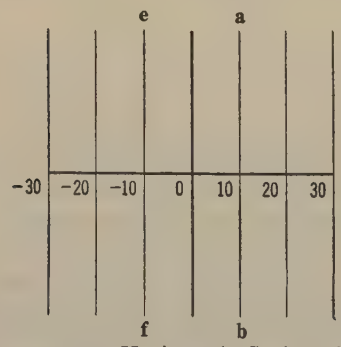


FIGURE 33.—Horizontal Scale with Series of Vertical Lines

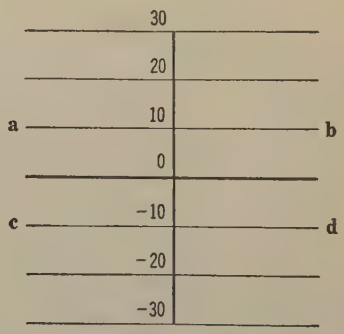


FIGURE 34.—Vertical Scale with Series of Horizontal Lines

tween these two lines are equal for any and every point on them—*i.e.*, any point on the line "ab" will be the same perpendicular distance from the line "cd" that the point "10" on the horizontal scale is from the point "0." Constructing perpendicular lines at

each point of graduation on the horizontal scale, we get a series of parallel lines running in grate-like manner across the horizontal scale as shown in Figure 33.

It will be noted therefore that any point on the line which passes through the starting point is "0" and that likewise the value of the point on the scale through which any perpendicular line passes will hold good for any point on the perpendicular line. Any point on the line "ab," for example, has a value of 10; any point on the line "ef" has a value of −10; and so forth. Hence it may be said that any value on the horizontal scale extends infinitely in both directions.

The same may be said for the vertical scale. Perpendicular lines drawn through the points of graduation on the vertical scale as shown in Figure 34 will likewise have the values respectively of the points on the scale through which they pass.

Coordinate Lines.—Perpendiculars drawn through the graduation points on two scales associated as in Figure 31 provide another device for graphic representation as shown in Figure 35 below. This device consists of two series of parallel lines arranged perpendicular to each other. Each series has the same characteristics as were noted above for the single scales. In addition, there are certain other important points that should be noted:

1. The device makes it possible to determine the value of any point in any quadrant in terms of a given scale. The horizontal scale shows how far the point is from the starting point horizontally; the vertical scale how far it is from the starting point vertically. The intersection of the lines drawn perpendicularly from these two points on the scales will determine the point in the field. The point "p" in Figure 35, below, for example, in the first quadrant has a value horizontally of "10" and vertically of "15."
2. From what has gone before, it is apparent that any value in the first quadrant is positive with respect to both the horizontal and vertical scales; any point in the second quadrant is negative with respect to the horizontal scale and positive with respect to the vertical scale; any point in the third quadrant is negative with respect to both scales; and any point in the fourth quadrant is positive with respect to the horizontal

scale and negative with respect to the vertical scale. This explains why the zero point is generally at the lower left-hand corner of the grid in practical work. We usually deal with positive values and hence use the first quadrant.

3. It will be recalled that the perpendicular lines are drawn only at the graduation points for convenience. Such lines may be drawn also at any other point on the scale. In practice we usually work from imaginary lines and do not actually construct them.

Summary.—By way of summary we shall indicate the terms that are generally used in connection with this fundamental ground-

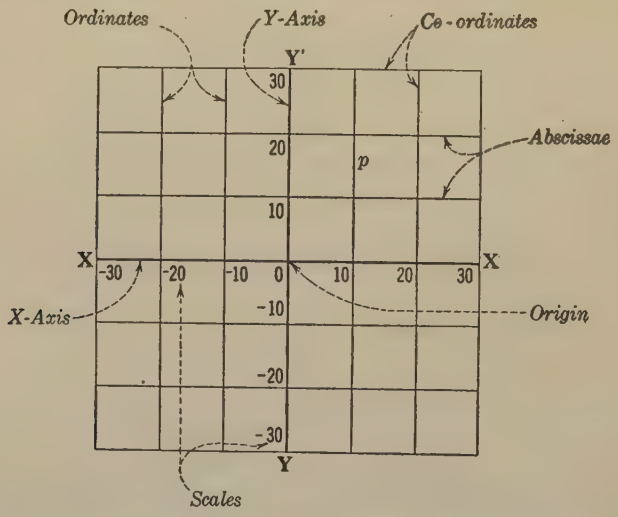


FIGURE 35.—Vertical and Horizontal Scales Associated, Each Having a Series of Parallel Lines
Terms used are also illustrated.

work of graphic representation. Figure 35 will help to clarify their meaning. Their significance in practical graphic work is developed more fully in the next chapter.

1. *Origin*—This is the point which is arbitrarily selected as the starting point. It is also called the zero point.
2. *Axes*—The two perpendicular lines which run through the origin—the one vertically and the other horizontally—are

the axes. The vertical line is called the Y-axis; the horizontal one the X-axis.

3. *Scale Numerals*—The figures placed along the axes at the points of graduation are called scale numerals. They indicate the values of the graduated points on the axes.
4. *Scales*—The graduated points on the axes together with the scale numerals are called the scales.
5. *Ordinates*—The vertical lines running perpendicularly to the X-axis through the several graduated points on this axis are called ordinates.
6. *Abscissae*—The horizontal lines running perpendicularly to the Y-axis through the several graduated points on this axis are called abscissae.
7. *Coordinates*—These are the two series of lines—the ordinates and abscissae—which run perpendicularly to each other. One or more pairs of these lines make up what are called coordinate lines. It is this type of ruling on graph paper which gives it the name "coordinate."
8. *Quadrants*—These are the four parts or sections in which the axes divide the field as indicated in Figure 31, page 276.

The idea underlying all chart construction is that whether the form of the graph be linear, area, volume, pictorial, or other, it will require some sort of scale—a unit of measurement will be selected to represent a unit of data. The essential elements of the scale enter into all types and forms of graphs. If the student thoroughly masters the significance of the foregoing, he will have little difficulty in grasping the method of constructing and using the statistical graph.

The Construction and Use of Graphs.—The questions of fundamental importance in graphics are the methods of constructing the various types and forms of graphs and the uses to which they may be put in business management. Consequently, the next few chapters will be devoted to the discussion of these two aspects of the subject. Consideration will be given to the fundamental elements of statistical graphs, the detailed steps to be followed in developing graphic charts, and the services which graphic charts render in the management of a business organization.

CHAPTER XXIX

THE BAR AND CURVE

The Two Chief Types.—The present chapter is a continuation of the discussion of the ground work of practical graphics. Here we turn to the detailed method of procedure in constructing the basic types of graphs. Since most of the graphic work in practical business statistics will be accomplished by the use of the *bar* and the *curve,* the method of constructing the simple forms of these two types, will be outlined in detail. The same fundamental principles will apply in the construction of the various forms or adaptations of these types of graphs and to a certain extent in the construction of other types.

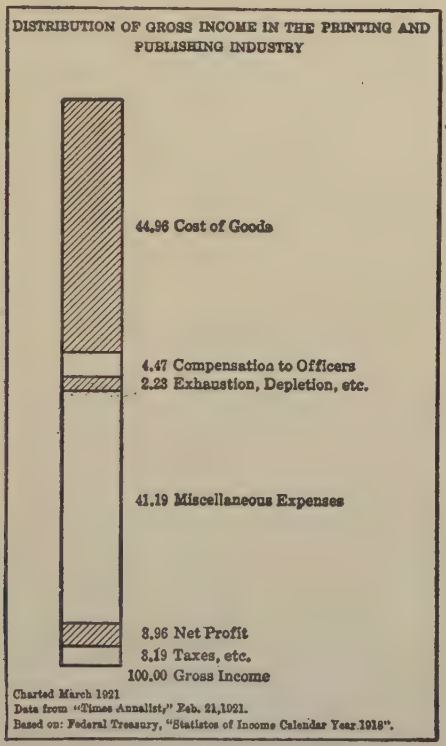
FIGURE 36.—Simple Bar representing Expenses of Chicago & Northwestern Railroad—August 1924
The bar may extend vertically in which case it is called a "column graph"

Column and Bar-Graph.—In this case the data are represented by bars, the lengths of the bars reflecting the magnitude of the data. Fundamentally, they are of two types—(1) the simple bar and (2) the component bar. The former is a single bar which represents a single figure as illustrated by Figure 36. The latter is also a single bar representing a single figure but it is divided into parts. The single figures represented by the several parts are components of the total figure which the bar as a whole represents. The component bar, therefore, may be said to represent a series of figures as shown in Figure 37. In either the simple or component type if the bars extend vertically we shall designate them as column graphs and if horizontally we shall call them bar graphs.

These are the basic elements of all bar-graph construction. More elaborate forms of the bar-graph are but combinations or

adaptations of these two basic elements. The more typical and widely used forms will be discussed in Chapter XXX, beginning on page 311.

Curve-Graph.—In this case, the data, strictly speaking, are represented by vertical bars, but practically their magnitude is


DISTRIBUTION OF GROSS INCOME IN THE PRINTING AND PUBLISHING INDUSTRY
44.96 Cost of Goods
4.47 Compensation to Officers
2.23 Exhaustion, Depletion, etc.
41.19 Miscellaneous Expenses
8.96 Net Profit
8.19 Taxes, etc.
100.00 Gross Income
Charted March 1921
Data from "Times Annalist," Feb. 21,1921.
Based on: Federal Treasury, "Statistics of Income Calendar Year 1918".

(From article by Arthur R. Burnet, *Administration*, Dec. 1921)

FIGURE 37.—Measuring the Printing and Publishing Industry by Yardstick of Tax Returns

The component bar.

designated by points determined by the intersection of coordinate lines, one extending horizontally and the other vertically, from points on the base lines which, according to the prescribed scales, represent the magnitude of the data. The data are really represented therefore by the distance from the base lines to the points. Fundamentally, they are likewise of two broad classes,

based on the character of the scales—(1) the *arithmetic* or *plain* and (2) the *geometric* or *ratio.*

In the *plain scale* the vertical rulings are equally spaced—*i.e.,* the distances between the lines which run horizontally across the chart field are equal. This is also true for the vertical lines. This scale or type of curve is illustrated by Figure 39, page 286. Other terms are also used to designate the *plain* scale such as "rectilinear," "difference," "amount-of-change" and "uniform."

In the *geometric* or *ratio scale* the spacings are graduated in accordance with the logarithmic scale similar to that of the slide rule. This may apply to either the vertical or horizontal scale or both. Figure 82, page 331 will serve as an illustration. It is also designated by other terms such as "logarithmic" and "rate-of-change."

There is a great variety of different curves which are used regularly in practical statistical work. They will be discussed in Chapter XXX, beginning on page 311.

Data Sheet.—The data sheet is a form used to set up the data conveniently for the draftsman and to furnish him directions. It contains the data which are to be graphed, set up in systematic form, the arrangement of the data taking the form which will most effectively facilitate the drafting work and at the same time reduce the probability of error to a minimum. To this end, a separate Data Sheet should be prepared for each chart or graph.

In addition to the data, the Data Sheet should carry at the top the exact title and sub-titles that are to appear on the chart. Likewise, it should also contain scale-captions, the source of the data, explanatory and footnotes to be placed on the chart, date of compilation, the initials of the person who did the compilation, and any other special or specific directions or explanations which may be important. Figure 38 represents a typical Data Sheet.

This is a very important and serviceable feature of graphic work. It not only facilitates the work of the draftsman but also lessens the chance of error in plotting. It also serves as a very convenient means of checking the drafting work.

Thinking the Graphic Problem Through.—The importance of thinking the graphic problem through cannot be too greatly

COMPARISON OF FACTORY PAYROLLS AND NUMBER OF EMPLOYES

By Months for 1923 and 1924

New York State

(Index Numbers with June, 1914, as 100. Source of data: The Industrial Bulletin, January 1925, Page 94)

Time Period	Payrolls (1)	Number of Employes (1)
a	b	c
1914	94	95
1915	101	98
1916	131	114
1917	156	120
1918	197	122
1919	213	114
1920	264	118
1921	188	92
1922	197	99
1923	238	110
1924		
January	235	106
February	236	107
March	241	108
April	229	104
May	219	100
June	210	97
July	201	93
August	204	94
September	217	97
October	215	98
November	216	98
December	222	99

(1) Payrolls represent actual payroll expenditures and not authorized rates. The number of employes includes all whose names appear on the payroll, both male and female, and office and shop.

Scale Caption —1. Wording—"Index Numbers."
2. Position —Associate with scale at both right and left.
Graph Caption—1. Wording—Column "b," "Amount of Payroll."
Column "c," "Number of Employes."
2. Position—Associate with the graph.
Form of Graph—Straight line curve.
Size of Chart—Overall 8 in. high, 6 in. wide; grid 5 in. high, 7 in. wide.
Scales—Abscissae and ordinates every ¼ in. Begin vertical scale with 90 at bottom and increase by 10; horizontal scale, a year at each ordinate.

Tabulated Jan. 20, 1925—C.C.B. Checked Jan. 20, 1925—L.D.M.

FIGURE 38.—Data Sheet

emphasized. The character of the data, the purpose of the presentation, and the use to which the graphic form is to be put should be carefully considered before the work is undertaken. In other words, the imagination must be brought into play and an effort made to visualize the proposed graph in its final form and to evaluate it in the light of the end to be served. We must definitely establish what the chart is to show—what comparisons, associations, or relationships are to be emphasized in the construction.

Furthermore, we must determine definitely the use to which the graph is to be put. If it is to be used in conference, it will need to be large in size or a number of copies will be required; in the latter case care will have to be taken to develop the original so that a number of copies may be produced from it. If it is to fit into a series of charts, it will generally conform in size and character to others of the series. If it is to be made a part of a report, a bulletin, or a book, it must be worked up so as conveniently to serve the purpose. All such things must be carefully noted before the actual development of the graph is undertaken. We must, in brief, think the graphic problem through.

Steps in the Construction.—It will be beneficial at this point to get a picture of the process of developing the statistical graph step by step from beginning to end. In outlining these steps, we shall follow the detailed process of constructing the simple historical curve and then indicate how the same principles enter into the construction of the bar-graph. The application of these principles to the construction of the frequency curve is developed in Chapter XXXV.

Preliminary to the actual construction of the statistical graph is the selection of the type and form of graph that is to be used —*i.e.*, whether it is to be a curve-, bar-, map-, circle-, area-, volume-, or picture-graph and what detailed form of these types will be required.

This is not always easy since different forms of graphs may generally be used. In making this selection it is usually advisable to construct rough drafts of those which seem most promising, to facilitate the final selection. It will be remembered that the char-

acter of the data and the purpose to be served are the primary factors in the selection. The prospective reader should also be taken into account.

In the construction of the simple historical curve the following principal steps will be observed:

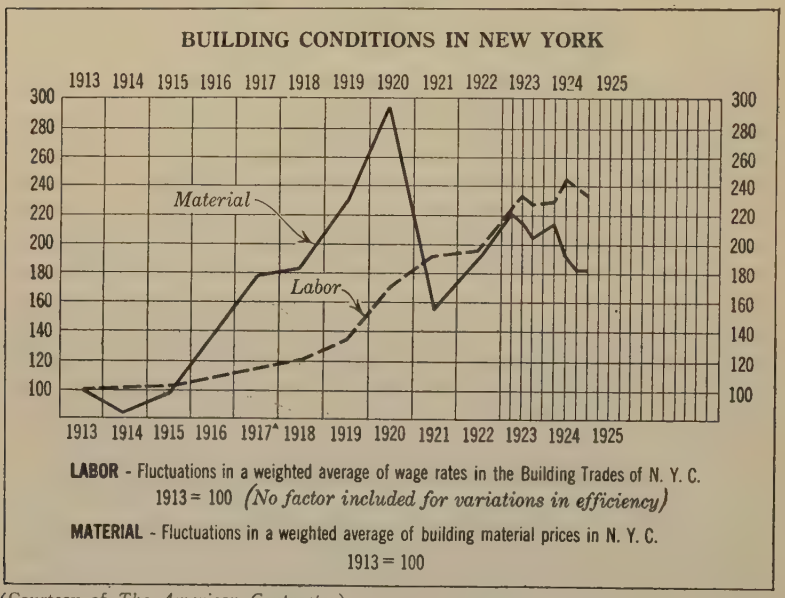
1. Select the grid or field.
2. Select the scales and enter the scale numerals.
3. Select and enter the scale captions.
4. Locate the points.
5. Connect the points.
6. Enter the title, explanatory notes, and footnotes.
7. Select and enter the graph designations.
8. Enter the date of issue, authority, and any other information that may be required.
9. Inspect for accuracy and completeness.
10. Arrange the data that are to accompany the graph.

Selecting the Grid or Field.—Primarily, the selection of the grid or field should not be an important question. The scale should be determined without regard for the grid and then the grid should be constructed to fit the scale. However, in practical work there are two conditions whose presence give the selection of the grid importance. We may wish to use a commercial paper, in which case the scales may have to be adjusted to the rulings of the paper, or we may wish the chart to be of certain fixed size, in which case also the scales will have to conform to the space allowed.

As a matter of practice in any case, the scale should be selected without regard for any specific grid or field. If necessary, it should then be adjusted to meet the requirements of a fixed field or size of chart. It should be emphasized that the effect of such adjustments on the final appearance of the graph should be carefully evaluated and in no case should an improper graph be presented merely for the sake of using a certain chart field or of keeping the chart to a certain size.

The Scales.—It will be remembered that the vertical scale carries positive numbers upward and negative numbers downward from the origin: the horizontal scale positive numbers to

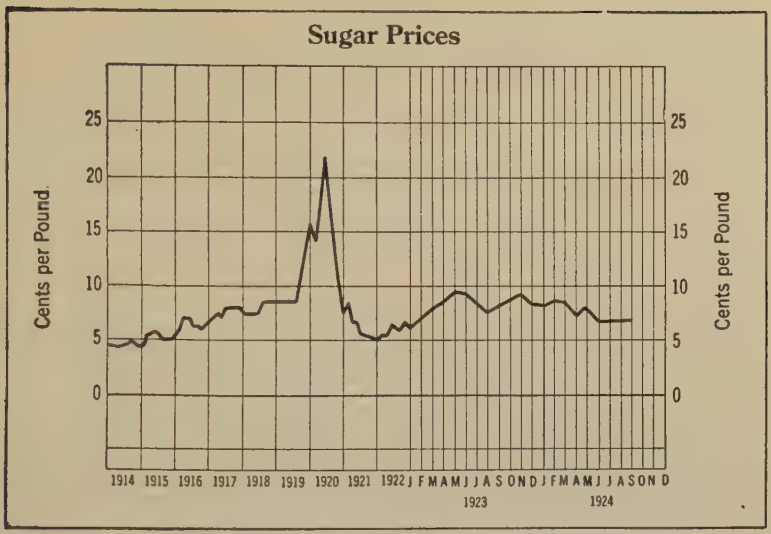
the right and negative numbers to the left from the origin. It will be observed also that positive values on both scales proceed in ascending order from the point of origin and negative values in descending order. Thus in Figure 39 the years on the horizontal scale are arranged in chronological order from left to right and the index numbers on the vertical scale are arranged in ascending order from the bottom upward.


BUILDING CONDITIONS IN NEW YORK
1913 1914 1915 1916 1917 1918 1919 1920 1921 1922 1923 1924 1925
300 280 260 240 220 200 180 160 140 120 100
Material
Labor
LABOR - Fluctuations in a weighted average of wage rates in the Building Trades of N. Y. C.
1913 = 100 (No factor included for variations in efficiency)
MATERIAL - Fluctuations in a weighted average of building material prices in N. Y. C.
1913 = 100

(Courtesy of *The American Contractor*)

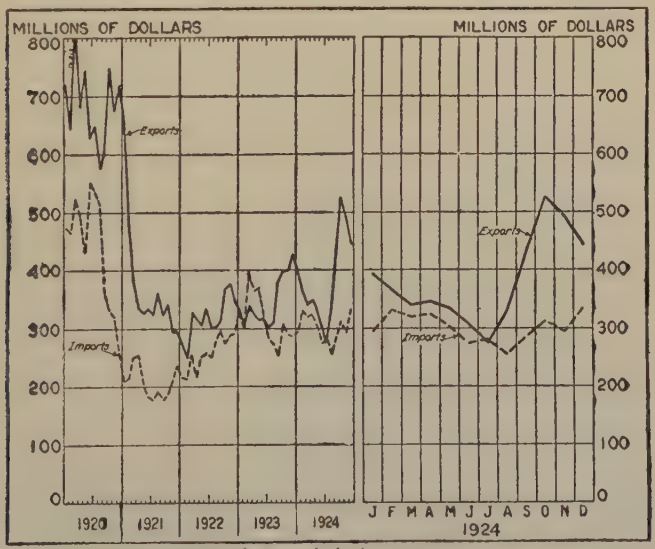
FIGURE 39.—Positive Values on Scales Proceed in Ascending Order from Point of Origin

In view of the fact that in practical work most of the data are positive, we use the first quadrant, as is pointed out on page 278. This leads to the practice of placing the vertical scale at the left of the chart field and the horizontal scale at the bottom as shown in Figure 39—*i.e.*, theoretically the axes are graduated to carry the scales and these sides of the field or grid correspond to the axes. This practice cannot be followed, of course, when negative data are being presented. Likewise when the field is large, or for any other reason, the scales may be repeated at the top and at the right side.



(Courtesy of *The Industrial Digest*)

FIGURE 40.—Unequal Space Given to Equal Time Periods.



(From *Commerce Reports*, by permission)

FIGURE 41.—Monthly Merchandise Exports and Imports

The right section repeats the data for 1924 on four times the horizontal scale of the left section; consequently the slopes are only one-fourth as deep. Good method of bringing out monthly figures.

Furthermore, it should be noted that equal distances at any place on the horizontal scale must represent equal values: likewise equal distances on the vertical scale must represent equal values regardless of the section of the scale at which the distances

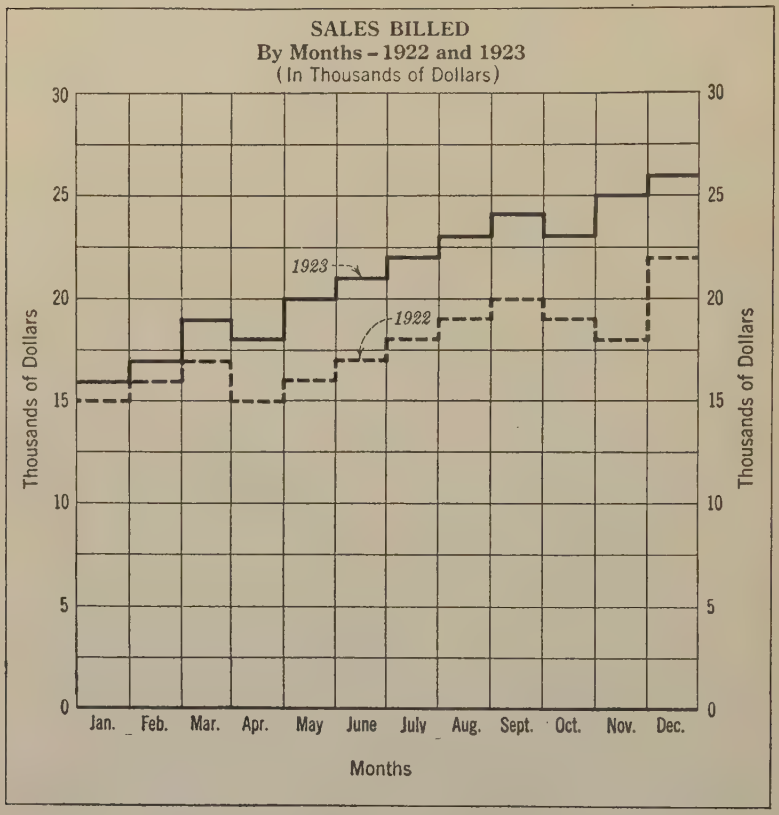


FIGURE 42.—Scale Unit Representing a Space
This scheme is not used extensively.

are taken. If the horizontal scale is to represent years, for example, equal distances must be given to the several years represented as shown in Figure 39, on page 286. Figure 40 illustrates a violation of this principle where a much greater distance is given to the last two years. Figure 41 shows a good method of bringing out monthly figures. It will be observed,

however, that a given distance on the horizontal scale may not represent the same value as an equal distance on the vertical scale.

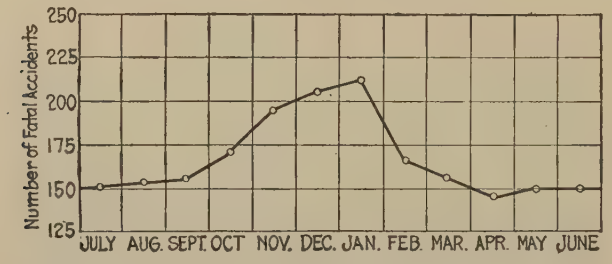
It should be kept in mind also that the purpose of the graph usually is to portray a general picture. Hence, there is generally no point in developing a finely graduated scale. Likewise, it is generally inadvisable to plot finely evaluated points: only the information which can be easily read should be recorded on the graph.

The Scale as a Space or Point.—The question is sometimes raised whether the data should be plotted on the ordinates or in the spaces. In view of the fact that the positions of the ordinates are more or less arbitrarily determined it is more fitting to inquire whether the scale unit should be a point as in Figure 39 on page 286 or a space as in Figure 42. The use of the latter is confined generally to continuous frequency data and to such facts as total payrolls, sales, and revenues in the historical field which extend unbrokenly over the period covered. Even in such cases, however, the comparative effectiveness of the resulting graph has established the practice generally of using a point rather than a space as the scale unit.

It should be emphasized that in either case equal spaces on the scale should be allowed for equal time periods as emphasized in last section. Even if the point be used, the distance between years, for example, should be twelve times the distance between months on the same scale. This is very important when plotting series with different time intervals on the same field as shown in Figure 39, page 286.

The Zero Point.—It is sometimes claimed that the absence of the zero point is apt to lead to a misinterpretation of the graph. The general impression from Figure 43, for example, it is contended, is that there were practically no accidents during the six months' period from April to August since these curves are so near the bottom of the grid. Figure 44, however, in which the zero point is shown reveals the true situation more clearly—that the number of accidents did not approach zero during the period indicated.

However, in Figure 39, on page 286, we are primarily interested in the base line of 100% with which we wish to compare the yearly fluctuations. The curve is of interest only in respect to the main line of reference—*i.e.*, the line which represents 1913. Here, we lose nothing by omitting the zero point.



(From article by C. B. Auel, *Management Engineering*, July 1921)

FIGURE 43.—Number of Fatal Accidents

Relation between the number of fatal accidents in industry and the months of the year. Note how the number of accidents increased during the winter months coincident with the short dark days, and fell off in the summer with the corresponding long light days. Zero point omitted.

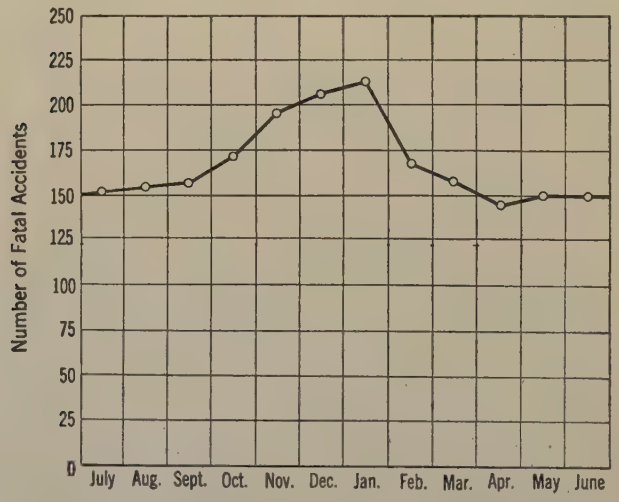


FIGURE 44.—Figure 43 Redrawn with Zero Line Shown

Fundamentally, therefore, the principle is broader than the one sometimes given—that the zero point should appear. The proper statement is *that the line of principal reference should appear.* If the "zero line" is the line of principal reference, it should be shown; if not, it may be omitted. The line of main reference may be the "zero line," a "normal," "100% line," "an average," or an arbitrarily selected point such as a quota in sales data. Whatever it is, as a regular practice it should not only be shown on the field but should be made to stand out by distinctive marking such as a heavier weight, or a distinctive color.

In cases where the graph is removed far from the zero or other line of main reference, we may realize economy of space by breaking the scale as shown in Figure 47. In practical work, however, there are cases where economy of space may require the omission of the line of principal reference in which cases the boundary of the grid should be a distinctly wavy line. In the frequency graph the zero of the horizontal scale should likewise be shown. An extensive discussion of this point is given by Earle Clarke, on page 662 of the Journal of the American Statistical Association for June 1917.

The Series and the Scales.—In making the choice as to which series of classes shall be placed on the abscissa and which on the ordinate, there is one fundamental principle that can be indicated—*the independent variable should be placed on the horizontal scale and the dependent variable on the vertical scale.* In the historical graph, for example, the time element is the independent variable and will be placed on the horizontal scale. In frequency data, the measurements which are the independent variable will be scaled horizontally and the frequencies which are the dependent variable will be scaled vertically.

It is apparent, therefore, that the ordinates need be constructed generally only where graduation points are registered on the horizontal base. This can be done since equal distances represent equal values and registrations on this scale will fall only at the points of the scale numerals. Hence, the points plotted will usually fall on a vertical ruling in the curve graph and the vertical

rulings will indicate the frequency of the plotting. This will apply only in case the grid or chart-field is specially constructed. If commercial coordinate paper is used the vertical lines will be fixed. But in such cases the ones which represent the values on the horizontal scale should be made heavier in order to bring out the graduations. When they are specially constructed, the ordinates should be of equal weight and only those entered which are required in the specific graph.

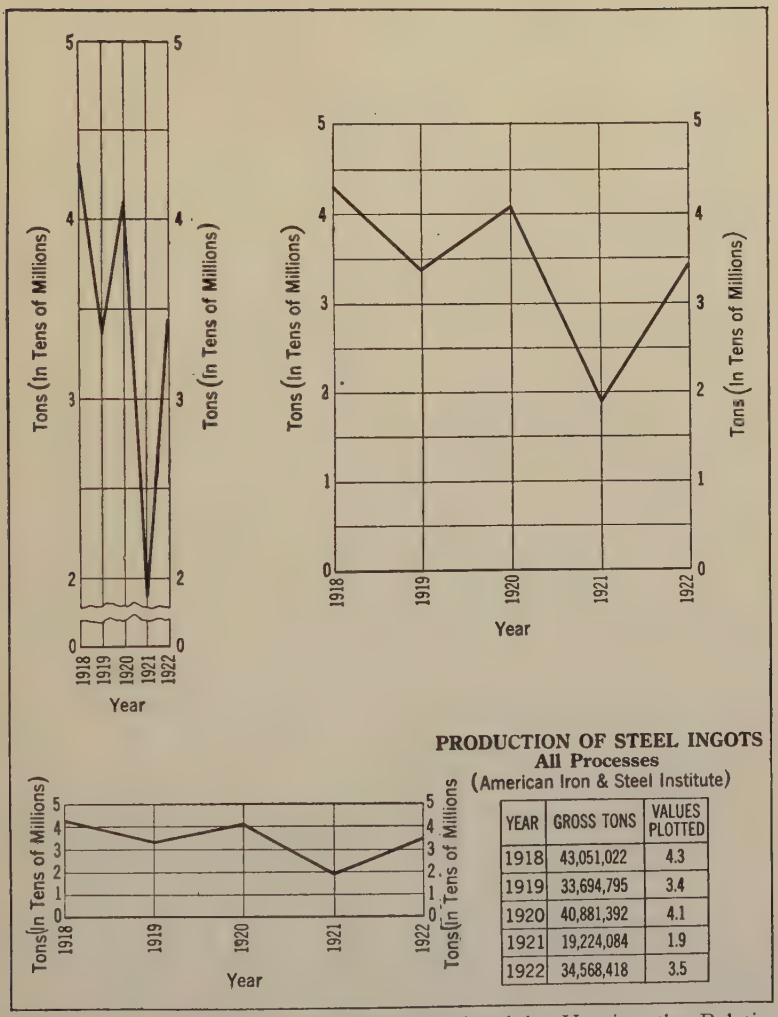
Ratio Between the Scales.—The matter of fundamental importance in the selection of the scale is the ratio between the vertical and horizontal scales as illustrated in Figure 45. These are made extreme in order to emphasize the point and while such extremes would hardly occur in practical work, less apparent discrepancies may occur and they may be quite enough to effect a misrepresentation. In reading graphs that others have constructed we should likewise pay particular attention to the relative value of the scales.

In determining just what this relation should be, we have to use our judgment to a great extent. No directions can be given which can be finely applied in all cases. There is however, a fundamental principle which will be helpful in practical work, as follows:

> The ratio between the two scales should be such that the smallest variation of importance in the data will be satisfactorily reflected. If we are presenting wage data, for example, on a weekly basis, and we consider a fluctuation of 50 cents or $1 a week as important, then the scales should be so selected that a fluctuation of 50 cents or $1 will be clearly brought out.

This point should be kept constantly in mind in selecting the scales. As a matter of practice, it is sometimes worth while to draw rough drafts of the curve with distinctly different ratios between the two scales—*i.e.*, one in which the vertical scale units are larger than the horizontal, another less than the horizontal, and a third equal to the horizontal. These may be varied to determine the effect on the graph.

Of course, if the graph is confined to a fixed field, which is



PRODUCTION OF STEEL INGOTS
All Processes
(American Iron & Steel Institute)

YEAR	GROSS TONS	VALUES PLOTTED
1918	43,051,022	4.3
1919	33,694,795	3.4
1920	40,881,392	4.1
1921	19,224,084	1.9
1922	34,568,418	3.5

FIGURE 45.—Illustration of the Effect Produced by Varying the Relative Values of the Vertical and Horizontal Scales
(The data are the same for the three figures)

relatively small in comparison with the range of the data the procedure is simplified. We simply compare the range of data in the series to be plotted—*i.e.*, the highest and lowest numbers—

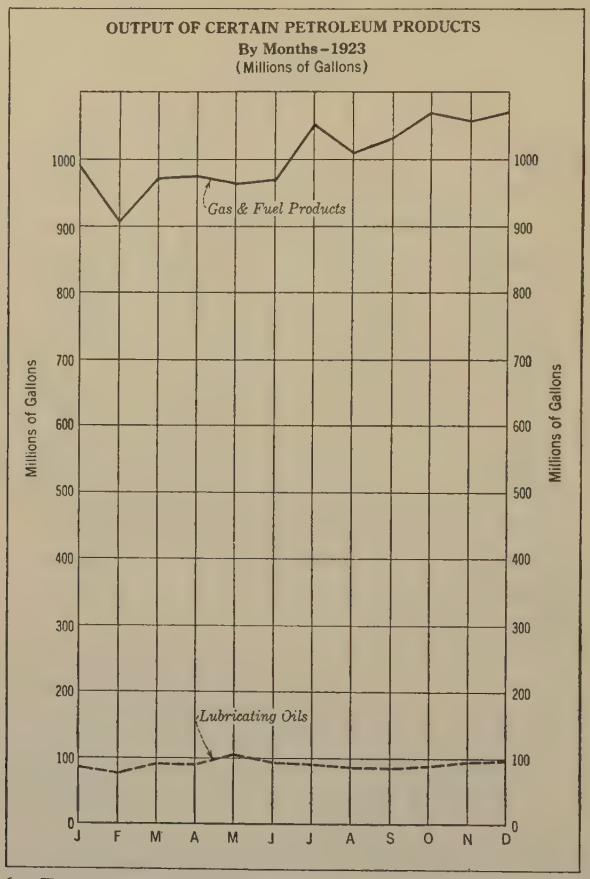


FIGURE 46.—For Purpose of Comparison these Curves Should be Brought Nearer Together

with the space available and select a scale unit which will permit the plotting of these extremes on the field. As was pointed out on page 285, however, the general character of the graph should never be seriously altered in order to fit it to a certain field. In any case, we should examine the graph to determine whether its

essential features correspond to the facts as indicated by the tabular data—*i.e.*, we should make sure that the impression given by the graph corresponds to the facts in the case.

In practical work in a business organization this problem is generally not so difficult as the rather lengthy discussion of it

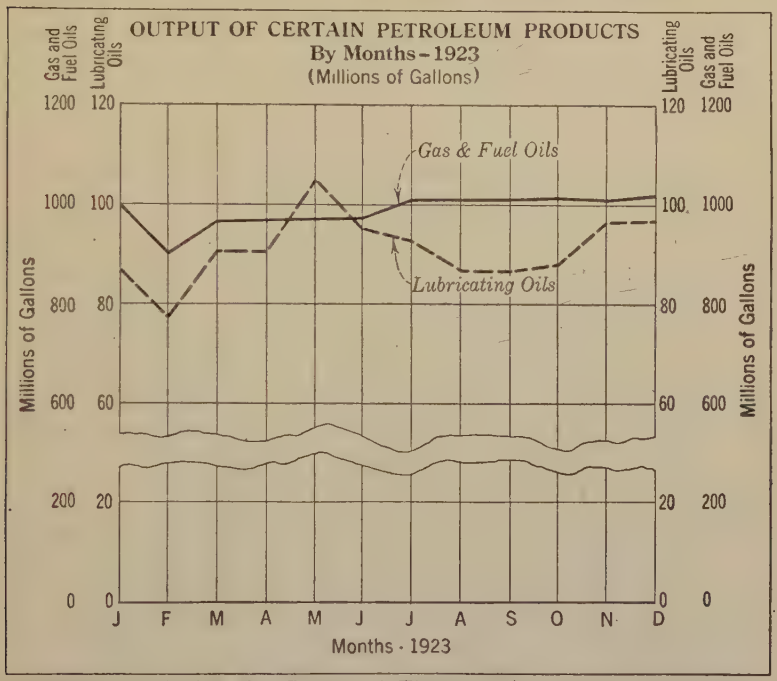
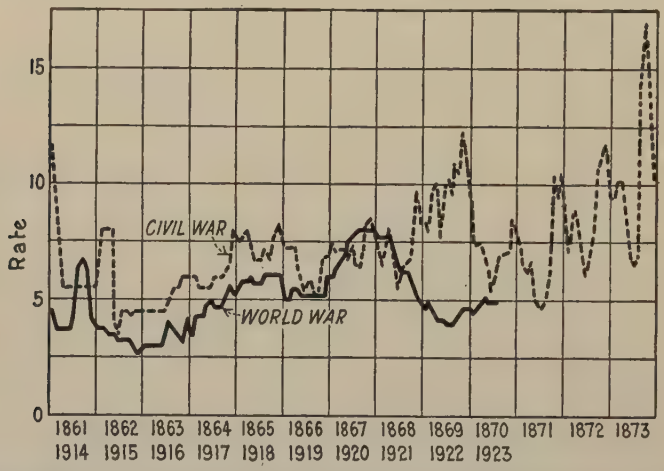


FIGURE 47.—Bringing the Curves of Figure 46 into Juxtaposition by the Use of Two Vertical Scales

here would indicate. In the regular statistical service we generally learn the relative scales to use for each type of data and, for the most part, the decision is made once for all.

Bringing Curves into Juxtaposition.—Different curves on a single field may be so widely separated as to make comparison difficult. Such curves may be brought into juxtaposition, (1) by omitting sections of the scale, or (2) by using two or more scales. These methods are illustrated by Figures 46-52. It is generally preferable to use the first of these methods since in the latter

case equal fluctuations in the curves do not reflect equal variations in the data and hence the curves are not comparable. If, however, the smallest fluctuation of importance is materially different in the two sets of data, we should use more than one scale. We should attempt in such cases to select the scales so that an equal fluctuation in the two curves would have the same significance. For example, if a fluctuation of $0.50 in one set of price data is equivalent to a fluctuation of $1 in another, the scales should be so selected that a fluctuation of $1 in the latter will show the same



(From article by W. Randolph Burgess, *Management and Administration*, Dec. 1923)

FIGURE 48.—Average Monthly Interest Rates on 60-90 Day Commercial Paper During the Civil War and World War Periods

Illustrates the method of bringing curves into juxtaposition by the use of two horizontal time scales.

change in the curve as a fluctuation of $0.50 will show in the former. This is much more satisfactory than the method sometimes suggested of making the two scales proportionate to the arithmetic averages of the two sets of figures to be plotted.

Curves may also be brought together by the use of relatives or percentages (Figures 51, 52) but these methods change the form of the data.

The Scale Caption.—The scale caption indicates the denomination of the units of the scale. It corresponds in a general way to the caption in the statistical table and likewise should be brief

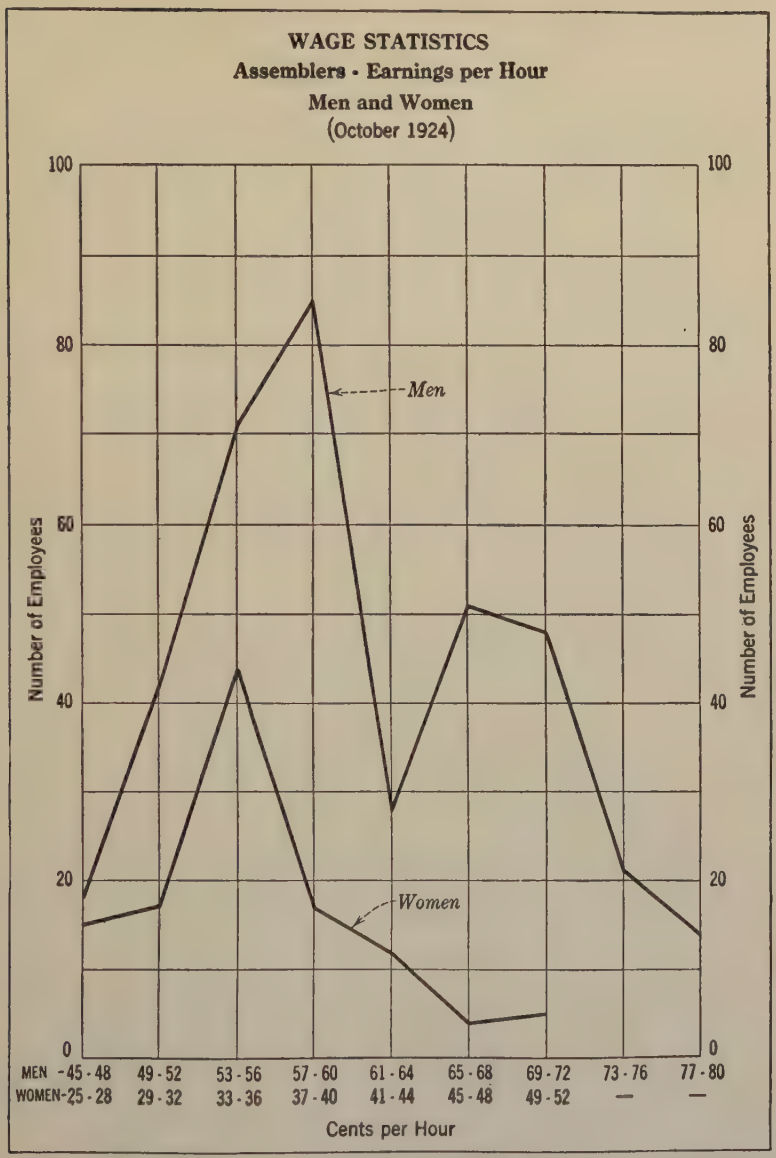

WAGE STATISTICS
Assemblers · Earnings per Hour
Men and Women
(October 1924)
100
80
60
40
20
0
Number of Employees
Men
Women
MEN -45 - 48 49 - 52 53 - 56 57 - 60 61 - 64 65 - 68 69 - 72 73 - 76 77 - 80
WOMEN-25 - 28 29 - 32 33 - 36 37 - 40 41 - 44 45 - 48 49 - 52 — —
Cents per Hour

FIGURE 49.—Bringing Curves into Juxtaposition by Means of Two Horizontal Frequency Scales

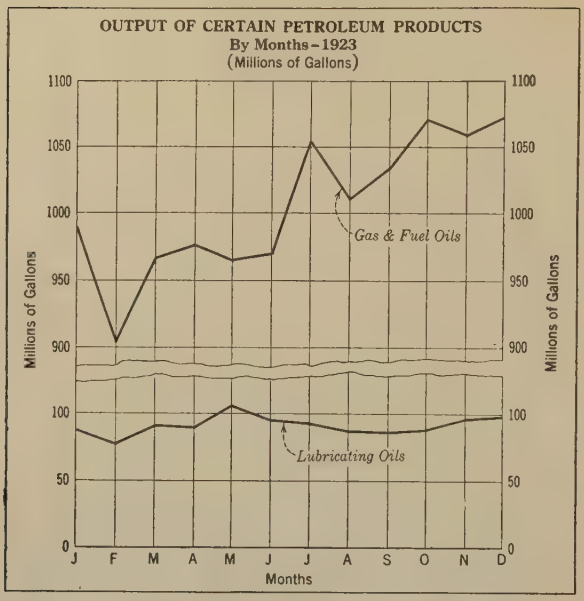


FIGURE 50.—Bringing the Curves of Figure 46 Together by Breaking the Scales

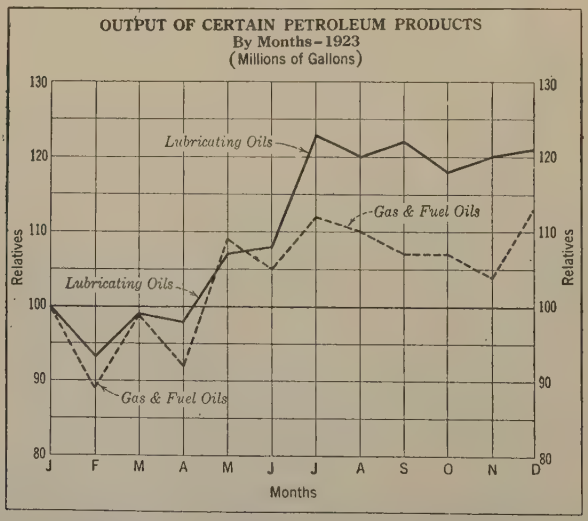


FIGURE 51.—Bringing the Curves of Figure 46 Together by Computing Relatives

and to the point. It is usually not difficult to formulate the caption. (See Figures 40, 42, and others in this chapter.)

The scale captions should be formulated and entered just as soon as the scales have been determined. This will keep constantly before the draftsman the scale denomination and will con-

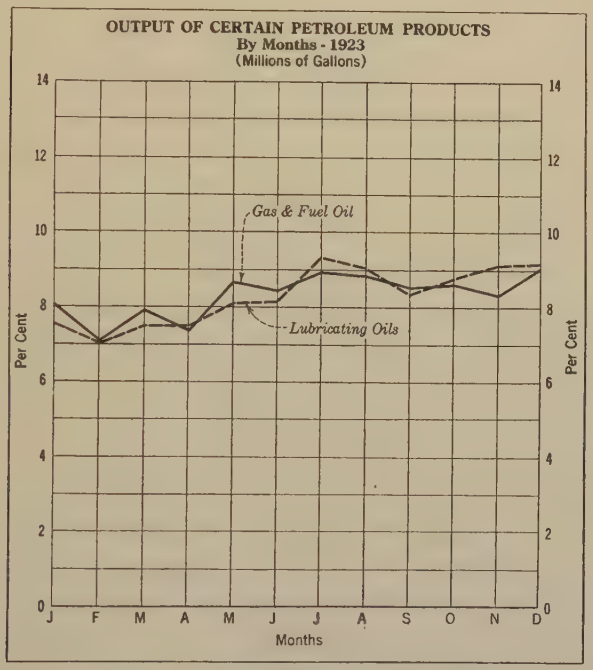

OUTPUT OF CERTAIN PETROLEUM PRODUCTS
By Months - 1923
(Millions of Gallons)
Gas & Fuel Oil
Lubricating Oils
Per Cent
Per Cent
Months
J F M A M J J A S O N D
0 2 4 6 8 10 12 14

FIGURE 52.—Bringing the Curves of Figure 46 Together by Expressing the Items as Percentages of Yearly Total

sequently tend to guard against error in plotting. If the scale is entered on both sides of the field, the scale caption should likewise be so entered. Furthermore, the entry of the scale caption should follow the general practice of having all entries on the chart so placed that they can be properly read from the bottom or right-hand side of the chart.

Locating the Points.—The points on the grid which are to serve as the basis of the graph are located by the intersection of

the coordinates. Thus the figure 3416 representing sales for April, 1923, is located in Figure 70, on page 322, by noting the month of April on the horizontal scale and 3416 on the vertical scale as indicated, and following the ordinate from the point representing April and the abscissa from the point representing the figure 3416 to the point where the two lines meet.

It should be noted that the ordinate and abscissa which determine any given point may or may not actually appear on the field. In practical work the ordinate will generally, though not always, be constructed for each value that is to be used on the horizontal scale. This will not be the case with the abscissa: it is only when the data coincide in value with the graduation points that the abscissa will appear. If the abscissa does not appear, the value of the data will be interpolated on the ordinate which represents the corresponding value on the horizontal scale. The point 87.5 in Figure 51, for example, falls just half-way between 85 and 90 on the vertical scale, 86.25 one-quarter of the distance, and so forth. Hence, the figure for February, 1923, is plotted on the February ordinate and 22 hundredths of the distance from 3500 to 3600 on the vertical scale.

Through the lack of data covering a particular period it may be necessary also to interpolate points for this period which will enable us to complete the graph. Such portion of the curve is usually distinguished by a different type of line, such as a dotted or dashed line when the line of the curve is solid. If we have no means of estimating the missing points, a straight line will be drawn between the two known points at the extremities of the gap. Frequently, however, we can estimate the missing data.

Connecting the Points.—It is the connection of the points located as above that gives definite shape to the curve-graph. The method may vary, producing three distinct forms:

1. *Line-and-angle Curve*—The method most commonly used is, beginning at the left, to draw a straight line from the first point to the second, from the second to the third, and so forth. This produces the "line-and-angle" or "peak-top" curve as illustrated in Figure 52 above. In frequency data it is the frequency polygon.

2. *Rectilinear or Stair-cased Curve*—If the scale designations on the horizontal scale represent spaces instead of points, the connection of the plotted points will produce a series of rectangles as suggested in Figure 42, page 288. The upper boundary of this figure designated by the heavy line is called the "rectilinear," "stair-cased," or "flat-top" curve.

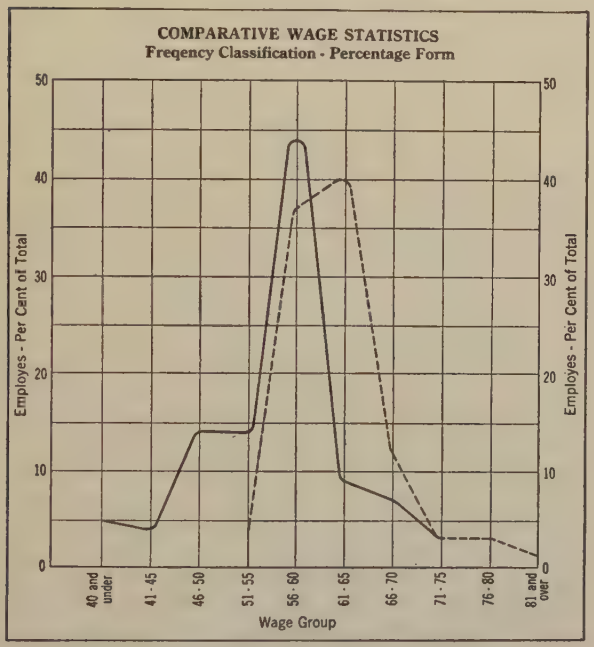


FIGURE 53.—The Rounded Frequency Curve

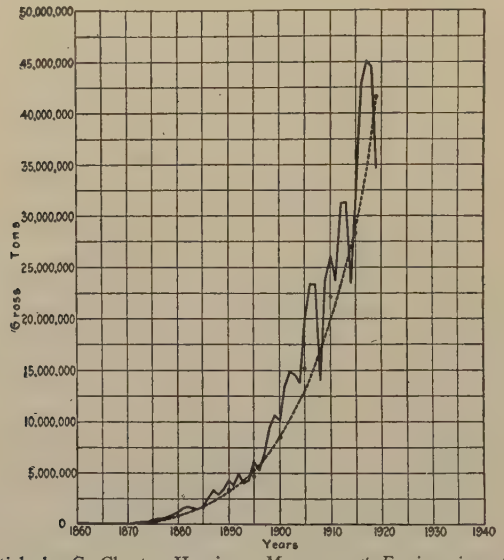
3. *The Smoothed Curve*—In speaking of the "smoothed" curve we will have in mind generally one of two things:

a. The line connecting the plotted points may be simply rounded out rather than constructed after the fashion of the line-and-angle curve. This will be accomplished by the free-hand method, with or without the aid of the French Curve or similar instrument. This curve will pass through all of the points just as does the line-and-angle curve. It is illustrated in Figure 53.

b. A curve may be drawn which takes the same general direction as the line-and-angle curve but which irons out its irregularities, giving as a result an approach to the ideal in a frequency group and the trend in an historical series. This curve will never pass through all points representing the actual data: it may or may not pass through some.

Methods of Smoothing Curves.—In an elementary treatise of this kind the general methods of smoothing curves noted in "b" above may, for practical purposes, be classed as:

1. *The Free-hand Method*—Whereby a smoothed curve is thrown directly over the plotted points in an attempt to smooth out



(From article by G. Charter Harrison, *Management Engineering*, Aug. 1921)

FIGURE 54.—The Production of Steel from 1860 to 1919
Smoothing curve by free-hand method

the irregularities. The position or course of this curve in any given case depends a great deal on the judgment of the person who draws it and, hence, there is no *one* curve of this form. Different persons will draw different curves. We try to visualize the ideal form of curve in a frequency series and the general trend in a time series and draw the curve in accordance therewith. Figure 54 will illustrate this method in the time series and 55 in the frequency series.

2. *The "Smoothed" Data Method*—This method involves the working out of a series of smoothed data from the actual figures by means of simple arithmetic computations. The smoothed data will be obtained by computing (1) a series of averages or (2) a series of totals. The computations may involve the simple average of the series, or the averages for the upper and lower half of the series, averages of quarters, or for smaller arbitrarily selected groups. The most frequently used computation in this group is the *progressive* or *moving average*. Likewise, most frequently used are the *progressive*

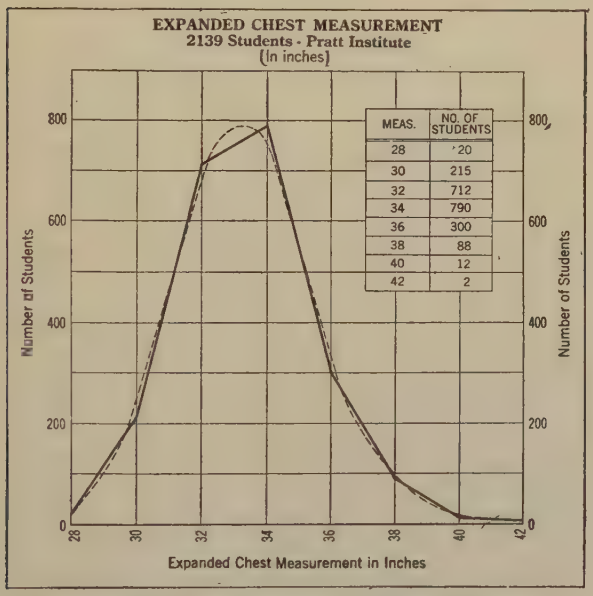
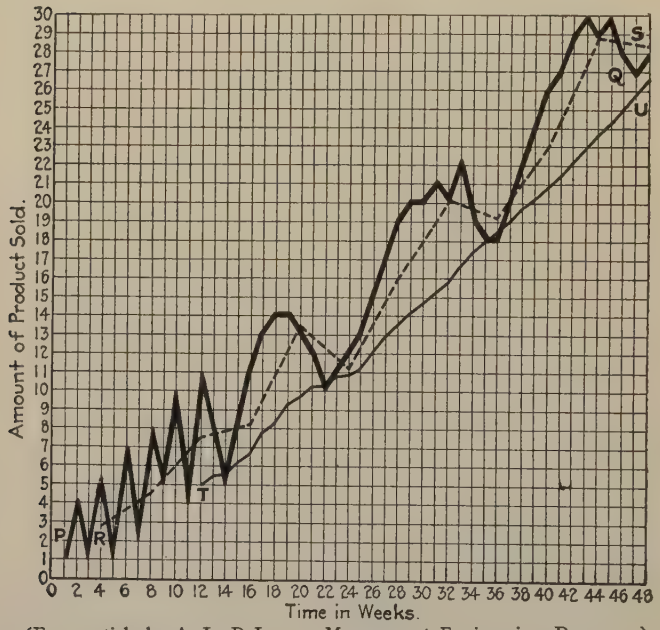


FIGURE 55.—The Smoothed Frequency Curve

or *moving total,* in which cases totals are taken of a certain selected number of items, the specific items included changing, however, progressively as explained in Chapter XXXIII. These methods for a simple time series are illustrated in Figure 56 and Figure 70 (page 322). While these methods are applicable to both the frequency and historical types of series, their principal use is in connection with the historical series.

3. *More Highly Mathematical Methods*—These methods consist generally of developing an algebraic equation of a curve

to fit the plotted points and representing it graphically. The method of least squares is perhaps used more extensively than other methods of this type: it is explained in Appendix E and its application is illustrated in Chapter XLIII. It is illustrated by Figure 57. Such methods, however, are employed comparatively little in elementary statistical work in the field of business and their discussion obviously falls outside the scope of a treatise of this kind. Detailed discussion will be found in the references given on page 351.



(From article by A. L. DeLeeuw, *Management Engineering*, Dec. 1921)

FIGURE 56.—A Chart of Sales from Week to Week

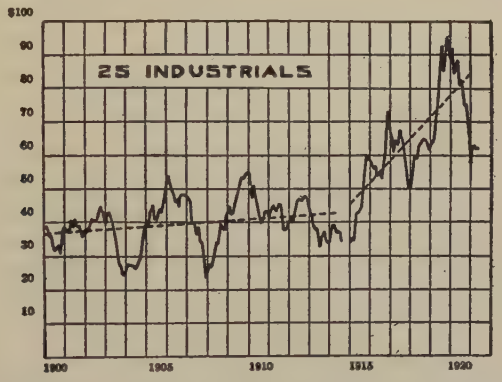
Curve PQ represents actual sales plotted weekly; RS actual sales plotted on a four week basis; and TU actual sales plotted on a twelve weeks' moving average basis.

It should be emphasized, however, that attempts to smooth data should not be made indiscriminately. We employ such methods for a specific purpose—*i.e.*, to get a picture of the true or ideal frequency curve or the trend of the time series—and we should utilize them only when they serve this purpose in a specific case. In other words, we are attempting to eliminate from the distribution irregularities resulting from errors in collection and classifi-

cation, from having only a relatively few data, and from any other cause apart from the nature of the data, and we should use the methods of smoothing outlined above only when they contribute to this end. Our aim in brief is to secure a better portrayal of the true situation than the limited number of data given us.

Title—Explanatory Notes—Footnotes.—These features of the graphic chart have been discussed in detail on pages 235 to 238. Their position on the chart is all that requires attention here.

Generally speaking, the preferable position for the title is at the top of the chart outside of the field; for the explanatory notes,



(From article by Leonard P. Ayres, *Administration*, Aug. 1921)

FIGURE 57.—Average Monthly Prices of 25 Industrial Common Stocks from 1900 to 1921

Dashed line, constructed by method of least squares, shows trends

at the top also, between the title and the field; and for the footnotes, at the bottom of the chart outside of the field. This is in accordance with the general principle of reading from the top of the page downward. We want the reader to see the title and explanatory notes before he inspects the graph and he will refer to the footnotes as he reads the graph. These positions also separate these parts of the chart entirely from the field or grid so that they do not detract from the effectiveness of the graph. They are illustrated by the figures given in this chapter.

Of course, they may be given other positions on the chart such as the bottom of the chart or on the field. The footnotes particularly may be boxed in and placed on an open section of the

field. The explanatory notes may be placed on separate memorandum if a lack of space on the chart requires it.

Graph Designations.—In presenting more than one series of data on a single field, means must be provided whereby the several curves may be associated respectively with the data they represent. This is generally done by using a different type of line for each curve and indicating by means of what we shall call "graph designations," sometimes also called labels, the series of data which each curve represents.

If the graph designation is not too elaborate and there is space on the field, it is usually preferable to place it right alongside the curve. This facilitates the reading of the curve by enabling us more readily to associate the graph with the data it represents. Figures 41 and 46 on pages 287 and 294 respectively, illustrate this arrangement.

However, if the presence of the graph designations on the field interferes with the effectiveness of the graph a separate "key" or "legend" should be used. In this case a small section of line similar to the curve is developed apart from the graph and the graph designation associated with it as shown in Figure 23, on page 267.

Final Inspection for Accuracy and Completeness.—Before the chart is finally released it should be carefully examined for accuracy and completeness by some one whose familiarity with the subject will enable him readily to detect errors. This examination should take the form of inspecting each of the foregoing steps one by one beginning with the title and following through until every phase of the chart has been inspected.

It is usually well to mark the chart definitely as the work of the statistical organization. This will enable the reader to locate the author for further inquiry. The date of issue should appear on the chart. The date the data are effective will appear in the title.

Data Should Accompany the Graph.—The foregoing steps finally complete the actual construction of the graph. There is one further point, however, which should be noted.

The data which the graph represents should generally accompany the graph. The presence of the data is important for several reasons. It gives the reader an opportunity to check the accuracy of the graph; it enables him to read the plotted points exactly which is frequently difficult if he has to depend on the scale alone; and it furnishes him with the data to use in different combinations or in other ways. All of these are important and the data should generally be present to accomplish them.

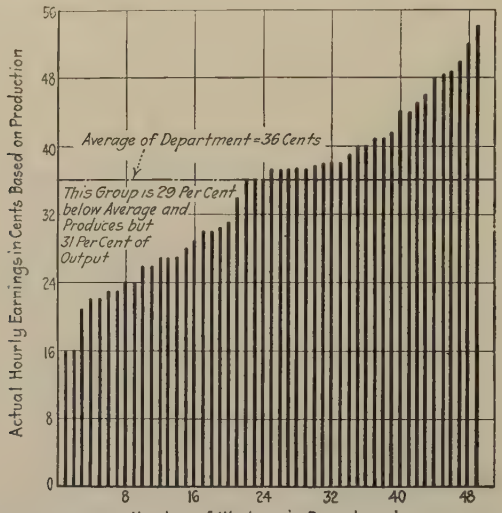
However, if the purpose of the graph is merely to present a general picture, such as may be the case in using graphs for advertising or publicity or in a conference discussion, the data will not be required. Furthermore, in practical business statistics frequently it will not be necessary to furnish the data since the detailed facts are always available in the statistical organization and can be obtained in a relatively short time. A further exception to the general rule is found in the use of popular graphics. Here it will be rarely advisable to accompany the graph with the data. In this type of graphic presentation we are not nearly so anxious to get the exact value of the data before the reader of the graph as we are to get the general impression which the picture gives. As a matter of practice in any type of graph, the facts of the specific case will determine whether the data should be presented.

There has been quite a bit of discussion as to the position the data should take when they do accompany the graph. The following suggestions are offered:

1. Listed on the same sheet as the chart but separated entirely from it.
2. On the chart but outside of the grid.
3. In a box in the grid.
4. Listed on a separate sheet from the chart.
5. At the points which represent them.
6. Listed along the top of the chart opposite the ends of the ordinates on which they are plotted.
7. Listed along the left edge of the chart opposite the abscissae on which they are plotted.

There is no established practice and adherence to any one of these positions must be based largely on an arbitrary decision. The

criterion of first importance is clearness of the graph and on this basis the first four in the list seem decidedly preferable. The argument for 6 and 7 is that data in the positions indicated facilitate the checking of the accuracy of the graph. But this is a matter of by no means first concern in the graphic presentation. Clarity and effectiveness of the graph are the fundamental considerations and the presence of the data in the positions of 6 and 7 certainly tend more to detract from than add to these considerations.



(Article by J. D. Johnson, *Management Engineering*, Aug. 1921)

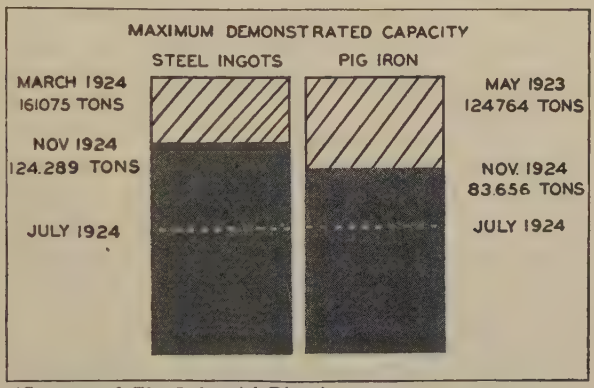
FIGURE 58.—Average Hourly Earnings of 49 Workers in a Typical Factory Department Engaged in a Bench Operation

Narrow bars in the bar-graph.

As a matter of practice we will place the data at different places at different times. If the data are few they may be placed in a box on the field or on the chart outside the field. If they are not too extensive it is very satisfactory to use half of the sheet for the data and the other half for the chart. We shall find cases in actual practice where the only feasible method is to furnish the data on a separate sheet. This will be true in a lot of the work in business statistics. In many cases, the graph is merely to accompany the tabular presentation to emphasize certain aspects of the data presented in tabular form, and in such event, the table furnishes the data to support the graph.

The Bar-Graph.—The same fundamental principles will be observed in the construction of the bar and column graph as have been outlined for the curve graph. An inspection of Figure 62, page 313 will substantiate this statement. Here, the years are placed on the horizontal scale and the series of facts on the vertical scale just as in the curve-graph. In fact, we may construct a curve to represent the data from these two scales just as they stand.

Generally speaking, the several steps outlined above in connection with the construction of the curve-graph will be followed also in the construction of the bar and column graph with the exception of one. Obviously, the method of "Connecting the Points" discussed on page 300 will not enter into this development. Here, the emphasis is placed on the distance of the points from the base while in the curve-graph it is the general direction of the curve and the fluctuations in it that are emphasized.



(Courtesy of *The Industrial Digest*)

FIGURE 59.—Wide Bars in the Bar Graph

There are three distinct points that must be kept in mind in the construction of the bar or column graph—(1) the width of the bar, (2) the distance between the bars, and (3) the scale to be used.

The width is relatively unimportant and will be a matter of judgment in any given case. The bars may vary from the thin line of Figure 58 (above) to the broad bar of Figure 59 (above), depending on the character of the data and the purpose to be served. It is important, however, that all of the bars which rep-

resent a given group or series of data in a single graph be of equal width.

The distance between the bars of a single group or series will likewise be decided on the basis of the facts of the given case. They may be placed against each other as in Figure 63, page 314, or they may be separated as in Figure 60 below. The question of space on the chart field will frequently be a determining factor. If two or more different classes of data are presented on the same graph, the bars will be grouped as in Figure 63, page 314, a relatively wide space being left between the groups. Lines may or may not be drawn across the field vertically to separate the groups.

When these points have been decided the bars or columns are simply constructed, reaching from the base to the plotted points. The proper scale is important also in the construction of the bar-graph. While the illustrations have shown columns, the graph may be constructed with bars extending horizontally from the vertical base. The references listed on page 351 give more extensive discussion of the bar-graph.

Conclusion.—When the chart has been finally released, a copy of it together with the data sheet should be filed conveniently in the statistical organization. Supplementing it also there should be filed the working papers, arranged systematically, of which the chart is the finished product. These records will not only serve as quick reference to facilitate checking any points in connection with the work but will also enable the statistical organization readily to answer questions which may come up or to furnish detailed facts bearing on the chart. The importance of routine reports and special studies as records is emphasized more fully on page 230.

In connection with the file copy, there should be noted the names of the persons who worked up the detail as well as those who supervised the work. Full and complete information relative to the source of data, methods of compilation, and so forth should also be developed. Putting the work papers in shape for final filing is by no means a small task and it is very important. They not only serve to support the particular chart but they also serve as sources of detailed data and computations. Much economy may be realized as a result of keeping them systematically at hand.

CHAPTER XXX

FORMS OF THE BAR AND CURVE GRAPH

Forms of Bar-Graphs.—The two basic elements of bar and column chart construction are the simple bar and the component bar. The various forms of the bar-graph presented here are combinations or adaptations of these two elements. Primarily the principal forms of the bar chart may be grouped into three broad classes—(1) series of simple bars, (2) series of component bars, and (3) combination of two or more series of simple or component bars. For convenience, the term "bar" will be used in the discussion to represent both bars and columns.

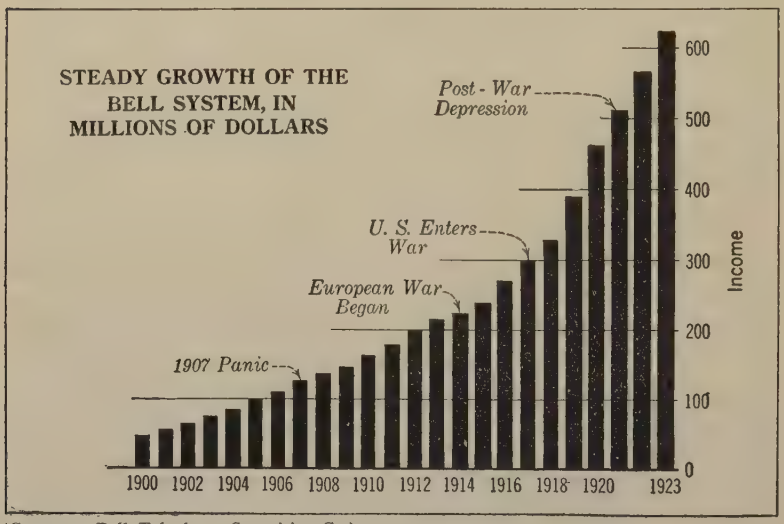
Series of Simple Bars.—In this case two or more simple bars representing related data are presented so that they may be readily compared. The bars may be extended horizontally or vertically. Figures 65 and 60 below will serve as illustrations.

This is a very simple and effective device for portraying comparisons of single figures whether personnel, financial, or operations. The different figures are represented by the lengths of the bars, which are projected from a common base. They may be used to represent graphically a great variety of data such as the number of employes, labor turnover, average rates of pay, expenses of different departments of the company, the purchasing production, and selling facts of subsidiary companies, and so forth.

Series of Component Bars.—This form of graph consists of two or more component bars systematically associated. It is sometimes called a component bar chart. Each bar represents a series of figures and the group of bars two or more series of figures.

There are two general classes of this form of graph, one in which all the bars are of equal length and the other in which the bars vary in length. The first is illustrated by Figure 61 below,

which is designed to show the total expenses of three railroad companies for August, 1924, broken up into their component parts, and expressed in the form of distributive percentages. It is sometimes referred to as the 100% compound bar-chart. Its most extensive use is in presenting data in the form of percentages. It may, however, be used to present absolute numbers as in the case of the distribution of workmen's time to different types of work, the distribution of a dollar of revenue, or any other cases where the total figures in the several series are equal.



(Courtesy Bell Telephone Securities Co.)

FIGURE 60.—Series of Simple Bars

The second form is typified by Figure 62 below. In this case the bars, which represent absolute numbers of varying magnitude, are unequal in length. This form of graph may be used to compare unequal absolute numbers and to break them up into their component parts, such as labor losses of the departments by causes or departmental expenses by classes.

It should be emphasized here that the component bar-graph is not effective for practical use. It affords ready comparison of only the series of parts which are adjacent to the base of the graph. The rest evidently cannot be well compared because they do not have a common starting point. Sometimes lines are drawn

connecting the boundaries of the component parts of the several bars, as in Figure 62 below: this improves the graph but does not overcome the disadvantage.

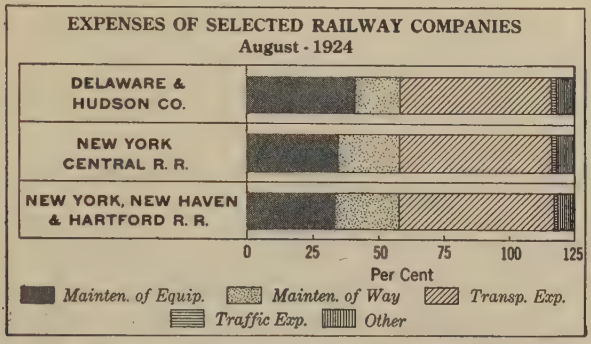
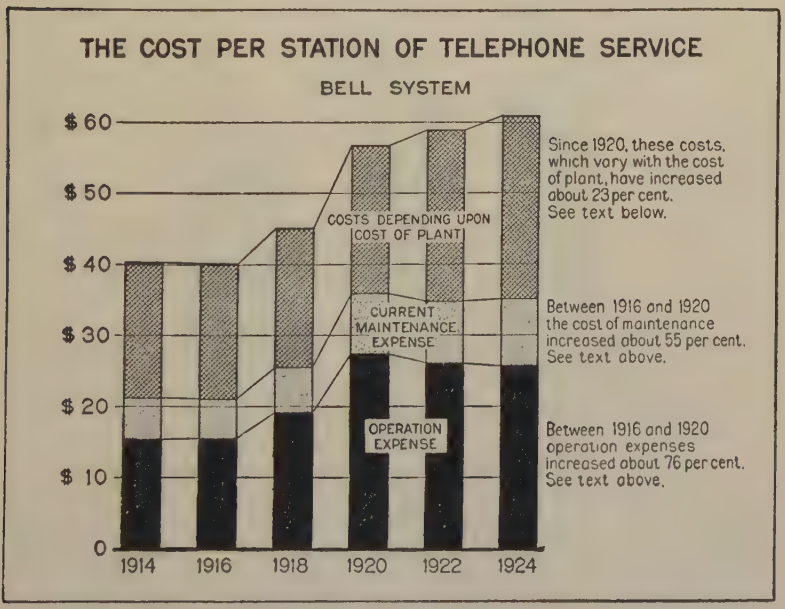


FIGURE 61.—Series of Component Bars—Bars of Equal Length


THE COST PER STATION OF TELEPHONE SERVICE
BELL SYSTEM
$ 60
$ 50
$ 40
$ 30
$ 20
$ 10
0
1914
1916
1918
1920
1922
1924
COSTS DEPENDING UPON COST OF PLANT
CURRENT MAINTENANCE EXPENSE
OPERATION EXPENSE
Since 1920, these costs, which vary with the cost of plant, have increased about 23 per cent. See text below.
Between 1916 and 1920 the cost of maintenance increased about 55 per cent. See text above.
Between 1916 and 1920 operation expenses increased about 76 per cent. See text above.

(Courtesy of The American Telephone and Telegraph Co.)

FIGURE 62.—Series of Component Bars—Bars of Unequal Length

Combination of Two or More Series of Simple Bars.—This is merely a combination of the form noted above. Two or

more series of data are represented by groups of simple bars and the several bars are associated so as best to facilitate comparison. Each bar represents a single figure and the different series of data are identified by distinctive markings. This form is sometimes called the multiple bar-graph.

In associating the several bars, those representing each series may be grouped together as in Figure 63 below, or the bars of

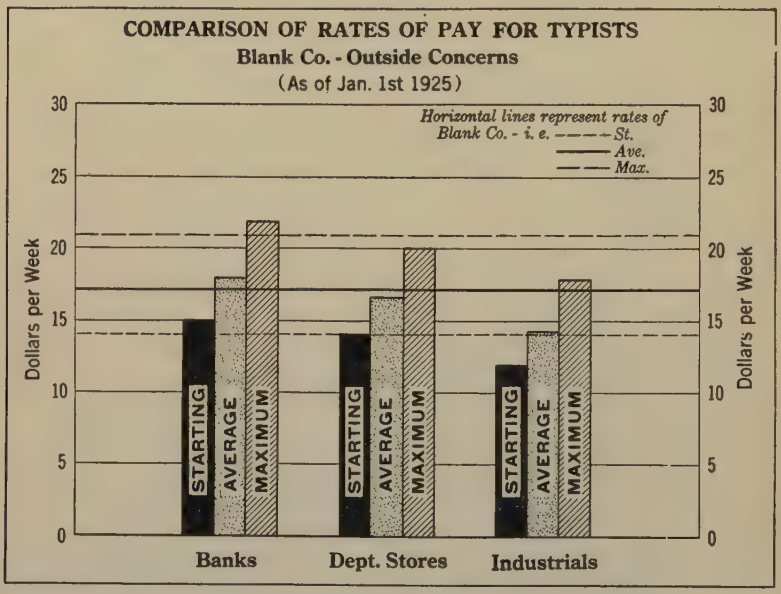


FIGURE 63.—Combination of Series of Bars with the Bars of the Separate Series Grouped Together
(Data are fictitious)

each series which are to be compared may be grouped together as in Figure 64 below. The particular method to follow in any given case will be determined on the basis of the facts of the particular situation. If the aim is primarily to compare the corresponding bars in the several series, the arrangement of Figure 64 should be used. If the principal comparison is between the several bars of the same series the arrangement of Figure 63 will be preferable.

Series of component bars described on page 311 may be combined in similar fashion.

Historical Data and the Bar-Graph.—The illustrations of the bar-graph given above have mostly represented data in cross-section or at a given time. While the principal use of the bar-graph is in representing this type of data, it will, however, be used in the graphic presentation of historical data also, (1) when the plotted points are few, (2) when we wish to make the graph stand out more prominently, and (3) when the purpose is merely

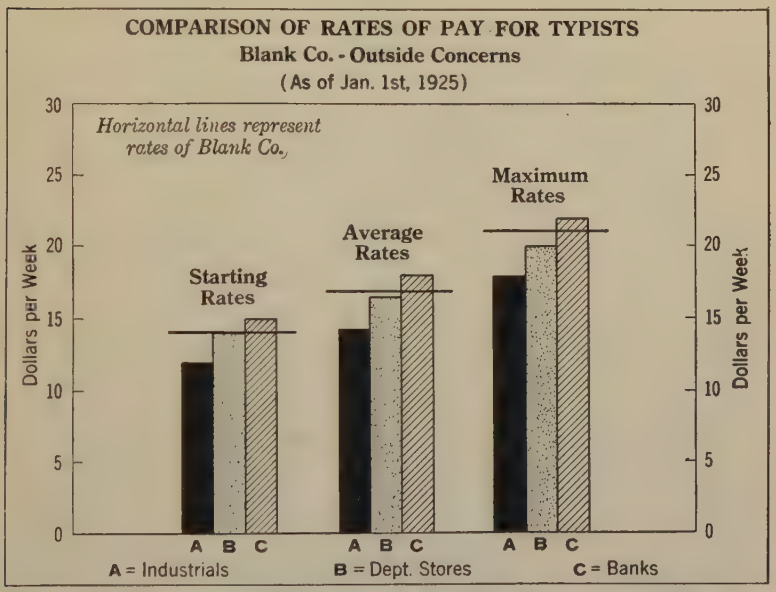
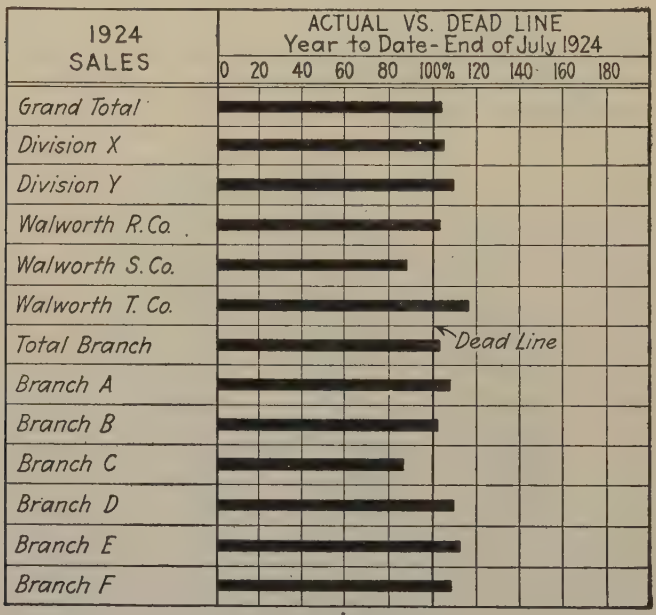


FIGURE 64.—Combination of Series of Bars with the Corresponding Bars of Each Series Grouped Together
(Data are fictitious)

to compare the facts of different time periods rather than to show the trend over a period of time. Figure 60 above is typical.

Miscellaneous Forms of Bar-Graph.—In addition to the foregoing more or less basic forms of the bar-graph there are many other variations which might be considered. We may note a few of the more typical ones.

Progress Charts.—Bar-graphs in the form of progress charts are used to represent a changing condition such as the output of a factory, the expenses, the revenues, and the progress of work in an organization. In such cases they are usually constructed to match with a standard or normal and are brought up to date daily, weekly, or as frequently as may be required in the particular case. This is merely a special application of the forms noted above; it is not a separate form of the bar-graph. Figure 65 will illustrate. Perhaps the most serviceable types of these graphs are mechanical ones obtained commercially.

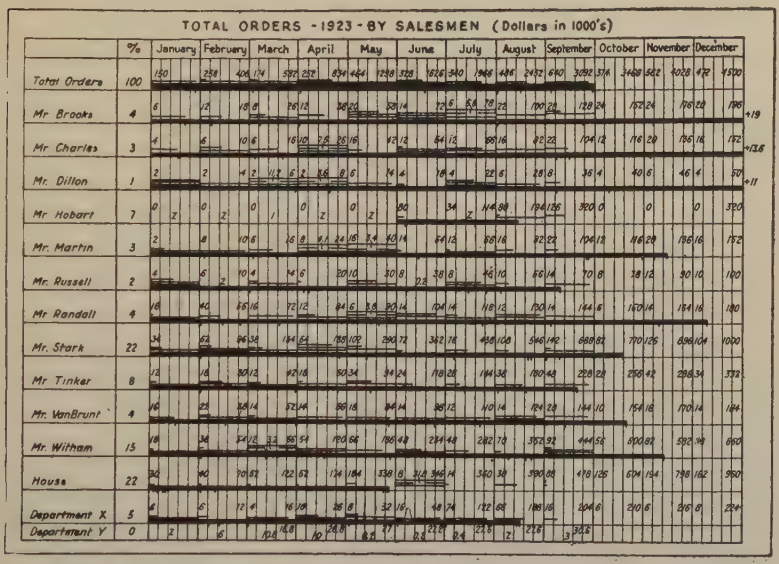


(From article by Joseph H. Barber, *Management and Administration*, Nov. 1924)

Figure 65.—Performance Comparison Against Forecast Sales
Progress bar-graph

Gantt Chart.—A special form of the progress chart is the Gantt progress chart illustrated by Figure 66 below. It is an estimate of future results and a record of actual results in comparison, both consecutively and cumulatively. The estimate for each time period—a month in Figure 66—is represented numerically by the figures entered in the left of the spaces and

graphically by the width of the space; the cumulative estimate is represented numerically by the figures entered at the right of the space and graphically by the total distance from the beginning of the first time period shown—January in Figure 66. The actual results are recorded only graphically, the cumulative results being represented by a relatively broad bar projecting from the left of the first time period shown and placed at the bottom of the horizontal space allotted to the item being represented. The consecutive results are represented by thin bars drawn across



(From Wallace Clark, *Management and Administration*, Feb. 1924)

FIGURE 66.—A Year's Sales Showing the Quantities Sold by Individual Salesmen

Gantt progress chart

the spaces allotted to the separate time periods, as many lines being drawn as the actual results are times the estimated.

Both the cumulative and consecutive bars represent percentages of actual results to the estimates or standards. Thus, in Figure 66 if the actual for January just equals the estimate a single thin line will be drawn across the space for January and the heavy bar will likewise be so constructed; if the actual is 150% of the estimate, one full line will be drawn across the

space just above the broad bar and another half-way across just above the first and the cumulative bar will be extended into the February space a distance equal to the percentage the excess for January is of the estimated for February; and so forth.

The Gantt progress chart is an exceptionally serviceable management device. It is applicable to a great variety of data in the personnel, financial, and operations fields, and should be freely used. More extensive discussion of its nature, construction, and use will be found in the references listed on page 351.

Duo-Bar Charts.—Sometimes two or more series of bars are presented on opposite sides of a common base, producing what is known as the duo-bar. These may be horizontal bars, the scale for one series extending to the right and that for the other extending to the left, or vertical, the scale for one extending upward and for the other, downward. This arrangement has serious disadvantages. It not only does not facilitate comparison of the two series but it also violates a fundamental principle that a scale which proceeds to the left or downward is negative. To represent negative numbers, it is of course proper but this is not the true duo-bar.

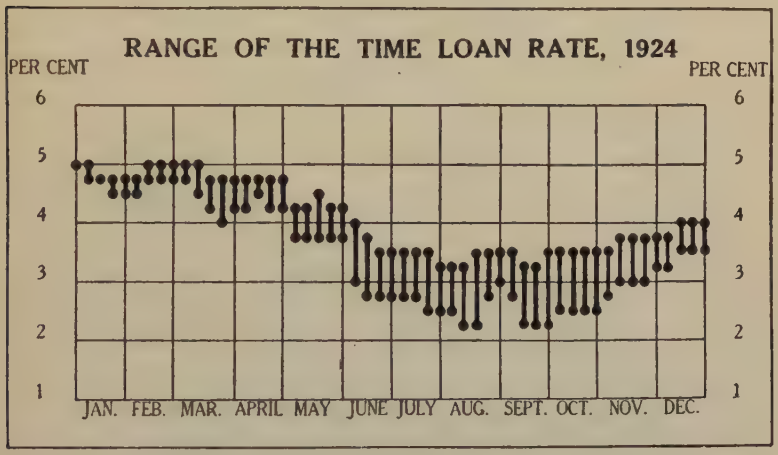
Other variations in the use of the bar-graph may be resorted to in special cases. More detailed discussion of these and other forms will be found in the list of references on page 351.

Forms of the Curve-Graph.—Both the arithmetic and logarithmic types of curves have many different forms. These forms are for the most part merely adaptations of the simple curve developed in Chapter XXIX, beginning on page 280. The student who is specially interested in the curve-graph will find more extensive development in the references listed on page 351.

Forms of the Arithmetic Curve.—Of the arithmetic type of curve there is just one form distinct from the simple curve or group of simple curves. This is the form in which the area described by one or more curves is given prominence. There are two principal kinds of this form of curve—the zone-curve and the band-curve.

The Zone-Curve.—The zone-curve represents series of data principally which have maxima and minima. The data repre-

senting the series of minima and those representing the series maxima are plotted and the space between the two is given a distinctive marking by means of coloring or otherwise cross-hatching. This type of curve is used freely in connection with security prices. The prices of the many individual sales range between rather narrow limits and are usually published showing the "high" and "low." Figure 67 below is a typical zone-curve. Other forms are constructed by, (1) plotting the two series as straight line curves and filling in between, (2) plotting the two series with a year represented by a space, (3) plotting



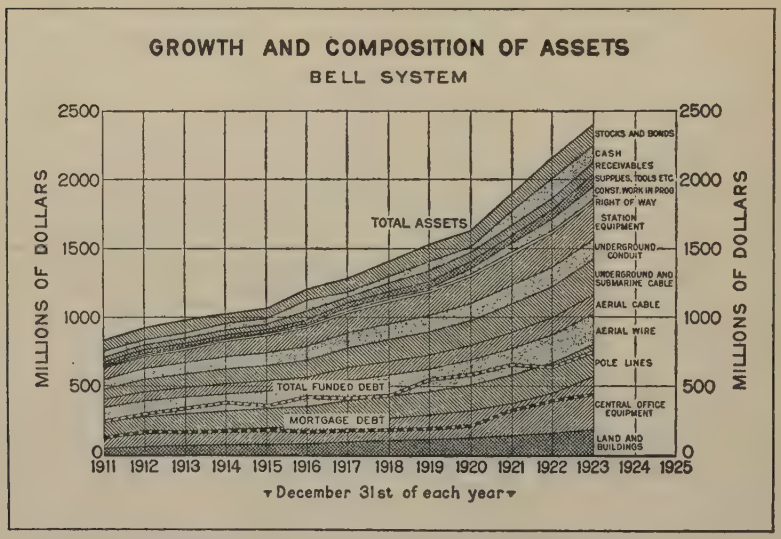
(From New York *Times Annalist*, by permission)

FIGURE 67.—Typical Zone Curve

the points and connecting each pair with a straight vertical line. The average may be shown with any of these forms by dots or circles plotted between the high and low points for each time period. Other kinds of data will require this form of graph such as the rates of pay asked by applicants for positions, numbers of units per individual sale, starting and full pay rates in the community, interest rates, commodity prices, and so forth.

The Band-Curve.—The band-curve is used to show the component parts of a whole. It is similar in general character to the component bar. In its construction the various component

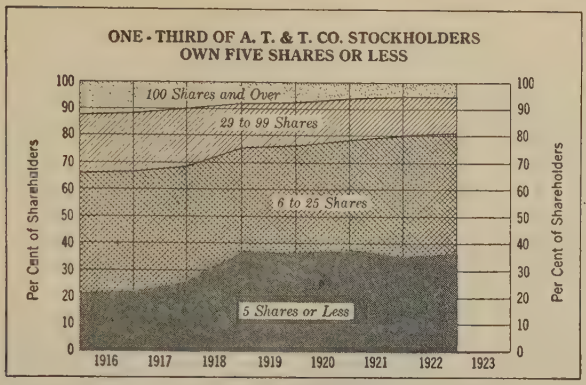
elements are brought out through the plotting of a series of curves, the spaces between which are given distinctive markings. The several curves do not represent the separate component parts as measured from the horizontal base but beginning at the base they represent the parts cumulatively upward until the last represents the total figure. In Figure 68 below, for example, the first curve from the base represents one component part, the second curve represents the first component part plus a second, and so on until the last curve represents a total of all of them. The spaces between the curves represent the component elements.



(From 1923 Annual Report of American Telephone and Telegraph Co.)

FIGURE 68.—Typical Band Chart—Absolute Numbers

As was true with the component bar, this form of curve may be expressed relatively or as 100% chart. In this form it is known as the 100% band or compound curve. The various component elements here are expressed in the form of percentages of the total and consequently the final curve will correspond to the 100% abscissa as shown in Figure 69 below. This form may likewise be used to present absolute figures providing they are equal such as the distribution or source of $100 of revenue.



(Courtesy of The American Telephone and Telegraph Co.)

FIGURE 69.—100% Band Chart

The Arithmetic Curve—Miscellaneous Forms.—In addition to the foregoing which are somewhat distinct in character, there are many other more or less distinct forms of arithmetic curves some of which take somewhat definite forms from the kinds of data they represent. The method of constructing them is quite similar to that of the simple curve. The following are typical of these curves:

1. The cumulative curve (Figure 70, curve marked "cumulative").
2. The moving total curve (Figure 70, curve marked "moving annual total").
3. The moving average curve (Figure 56, page 304, curve TU).
4. The Zee-curves (Figure 70).
5. The progress curve (Figure 71).
6. "Mountain" curve (Figure 72).
7. Divergence curve (Figure 73).

More extensive discussion of these miscellaneous forms will be found in the references listed on page 351.

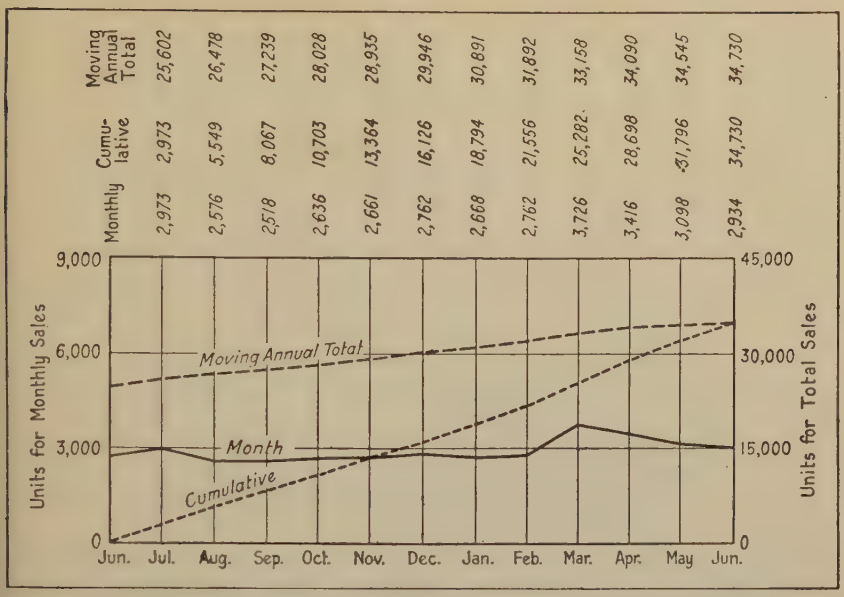
The Logarithmic Curve.—The fundamental principles of the logarithmic curve are developed in Chapter XXXI beginning on page 326. This curve is applied chiefly to historical data but may be used also to represent simple cumulative frequency data. The principles of constructing the graph in these different cases are

not different from those outlined in the previous chapter. Figure 83, page 335, will serve as an illustration.

Summary—List of Forms of Graphic Devices.—By way of summary there is presented at this point an outline of the principal forms of graphs. Those listed in groups III-VIII will be developed in subsequent chapters.

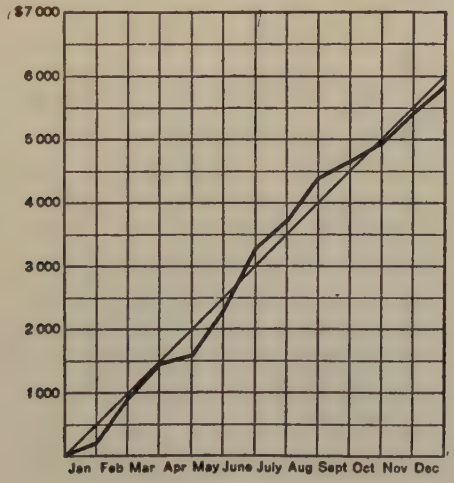
I. Bar-Graph
  A. The simple bar
    1. Single bar
    2. Series of simple bars
    3. Combination of two or more series of simple bars
  B. The component bar
    1. Single bar
    2. Series of component bars
    3. Combination of two or more series of component bars
  C. Miscellaneous forms
    1. Progress bar chart
    2. Duo-bar chart
    3. Gantt chart

II. Curve-Graph
  A. Arithmetic curve
    1. Historical
      a. Simple curve
      b. Zone-curve
      c. Band-curve
      d. Cumulative curve
      e. Moving total curve
      f. Moving average curve
      g. Zee-curve
      h. Progress curve
    2. Frequency—
      a. Simple curve
      b. Band-curve
      c. Cumulative or ogive
      d. Moving average curve
  B. Logarithmic curve—In practical work the use of the logarithmic or ratio scale is confined primarily to the simple and cumulative historical curves.



(From article by David B. Porter, *Management and Administration*, Jan. 1924)

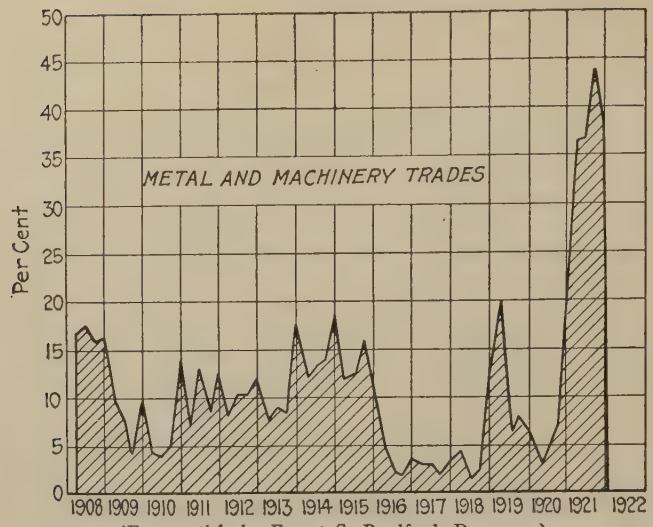
FIGURE 70.—A Chart of Sales Ending June 1923 (The Zee Curves)



(From article by Louis C. West, *Administration*, Feb. 1923)

FIGURE 71.—An Expense Chart

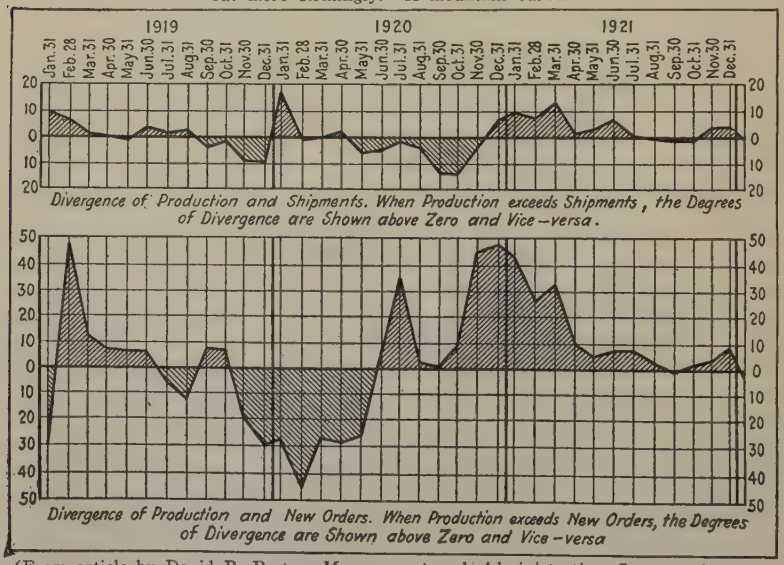
The light line represents the quota allowance and the heavy line the actual expenses. A progress curve chart



(From article by Ernest S. Bradford, Dec. 1922)

FIGURE 72.—Percentage of Organized Workers in the Metal and Machinery Trades Unemployed in Massachusetts, 1908-1921, by Quarters

It is obvious that this graph is constructed merely by filling in the space between the straight line curve and the horizontal base. Its purpose is to make the graph stand out more strikingly. A mountain curve.


1919
1920
1921
Jan.31 Feb.28 Mar.31 Apr.30 May31 Jun.30 Jul.31 Aug.31 Sep.30 Oct.31 Nov30 Dec.31 Jan.31 Feb.28 Mar.31 Apr.30 May31 Jun.30 Jul.31 Aug.31 Sep.30 Oct.31 Nov30 Dec.31 Jan.31 Feb.28 Mar.31 Apr.30 May31 Jun.30 Jul.31 Aug.31 Sep.30 Oct.31 Nov30 Dec.31
20 10 0 10 20
Divergence of Production and Shipments. When Production exceeds Shipments, the Degrees of Divergence are Shown above Zero and Vice–versa.
50 40 30 20 10 0 10 20 30 40 50
Divergence of Production and New Orders. When Production exceeds New Orders, the Degrees of Divergence are Shown above Zero and Vice–versa

(From article by David B. Porter, *Management and Administration*, Jan. 1924)

FIGURE 73.—Divergence Chart of Production and Shipments, and Production and New Orders

III. Map-Graph
  A. Shaded or cross-hatched maps
  B. Dot maps
    1. Data represented by size of dot
    2. Data represented by shading or cross-hatching the dot
    3. Data represented by frequency of dots
  C. Pin and ribbon maps
IV. Circle or Pie Graph
V. Area Graph
VI. Volume Graph
VII. Picture Graph
VIII. Combination Graph

CHAPTER XXXI

THE LOGARITHMIC OR RATIO GRAPH

The Construction of the Ratio Scale.—In practical statistical work in a business organization, logarithmic paper which can be obtained commercially will generally be adequate for most purposes and hence, there will be little necessity for constructing the

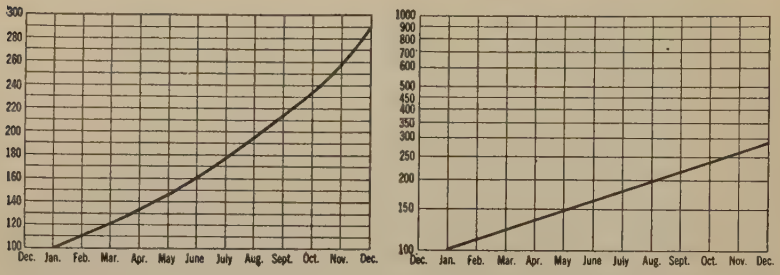


FIGURE 74.—Series Increasing by a Constant Ratio Take the Form of a Straight Line on the Logarithmic Grid and a Curved Line on the Arithmetic Grid

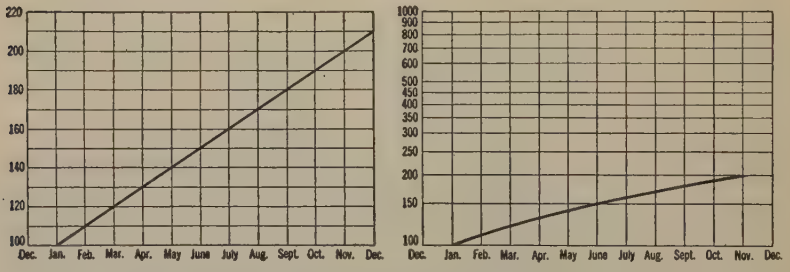


FIGURE 75.—Series Increasing by a Constant Difference Take the Form of a Straight Line on the Arithmetic Grid and a Curve on the Logarithmic Grid

scale. However, familiarity with the way such a scale is constructed will be helpful in the understanding and use of the logarithmic graph. In this type of scale, the distance between any two numbers on the scale represents the ratio of the one number to the

other rather than their difference, as illustrated in Figures 74 and 75 above.

Perhaps the simplest method of constructing the logarithmic scale is by means of the slide rule. In this case the graduations on the slide rule are transferred directly to the paper, the resulting scale being the same length as the slide rule—an exact reproduction of its scale. Scales of different lengths may be obtained by using slide rules of different lengths: or, by means of the device illustrated in Figure 76 below, a line of any length may be properly graduated directly from the slide rule.

FIGURE 76.—A Device for Graduating a Given Line in Accordance with the Scale of Another Line

The line "a b" is graduated in proportion to the scale on "a c" by drawing "a c" at any angle, connecting "c" with "b" and drawing lines from other points on "a c" parallel to "b c."

A second method of determining the logarithmic graduations on a given line is to assume an arithmetic or plain scale on the line from 0 to 1, Figure 77 below, and then plot on this scale the logarithms of the numbers 1 to 10 given in the accompanying table.

NUMBER	LOGARITHM
1	0.00000
2	0.30103
3	0.47712
4	0.60206
5	0.69897
6	0.77815
7	0.84510
8	0.90309
9	0.95424
10	1.00000

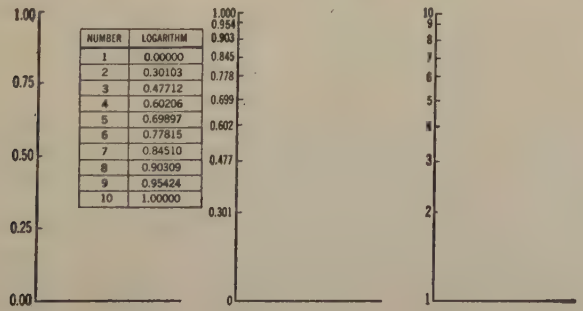


FIG. 77 FIG. 78 FIG. 79

This will result in the scale shown in Figure 78. By substituting the absolute numbers for their corresponding logarithms we have the final scale shown in Figure 79.

Finally, the procedure which most effectively portrays its nature is suggested by Professor Irving Fisher in an article appearing on pages 577-600 of the Quarterly Publications of the American Statistical Association for June 1917. Two principal steps are indicated in this method:

1. Label the horizontal lines of an arithmetic grid with numbers increasing in a given ratio as shown in Figure 80.
2. Since these odd numbers make plotting difficult, interpolate round numbers as shown in Figure 81 and finally eliminate the equi-distant lines.

It is clear from this explanation that we can shift the values of the scale designations leaving the spacings equi-distant as in the plain ruling or we can shift the spacings and let the numbers or scale designations remain as in the arithmetic or plain ruling.

Forms of Logarithmic Grids.—Two primary forms of logarithmic grids may be distinguished:

1. The semi-logarithmic in which one of the scales is logarithmically graduated and the other arithmetically graduated. In practical work the vertical or Y-scale is usually logarithmic and the horizontal or X-scale arithmetic. The term "arith-log" is used commercially to indicate this form of paper.
2. The *full logarithmic* in which both the vertical and horizontal scales are graduated logarithmically.

In addition to these, the "part-deck" form which shows only a part of the cycle, and the "split-deck" form which is made up of remote parts of two separate cycles, may be convenient in some cases but generally little advantage is gained by using them.

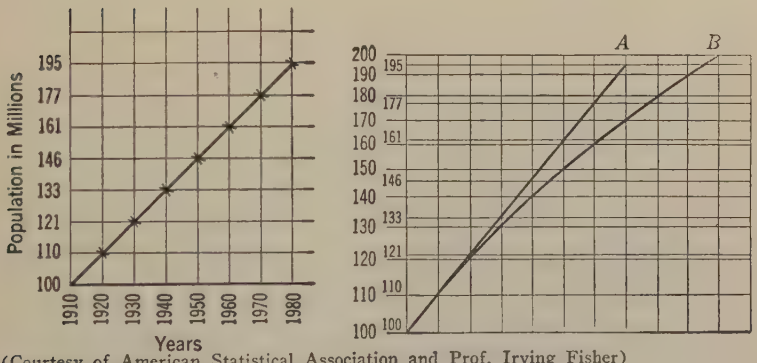
To facilitate the discussion, further consideration of the subject will be confined for the most part to the semi-logarithmic form in which only the vertical scale is logarithmically graduated.

Distinctive Features of the Logarithmic Scale.—From the practical point of view, the logarithmic scale differs from the arithmetic in certain definite and important respects. Chief among these are the following:

1. The vertical distance between any two points on the logarithmic scale reflects the ratio or percentage difference between the values representd by the two points. In other words, it reflects the percentage increase of the higher over the lower number and the percentage decrease of the lower from the higher. Likewise, the direction of the line connecting points on the logarithmic grid reflects the rate of change in the series of numbers plotted.

   These are the fundamentally important elements of the logarithmic scale. The student will find it helpful carefully to contrast this feature of the logarithmic scale with the arithmetic scale in which the distance between any two points reflects the numerical difference between the values represented by them.

2. As a corollary to No. 1, it follows that equal vertical distances on any part of the scale represent equal ratios even though the absolute numbers involved may be greatly different. Likewise, different portions of one or more curves, which take the same direction, reflect equal rates of change.



(Courtesy of American Statistical Association and Prof. Irving Fisher)

(80) (81)

FIGURE 80.—Logarithmic or Ratio Chart Made from Ordinary Ruled Paper by Labeling the Horizontal Lines for Ratios

The curve represents the future hypothetical population of the United States, assuming 10% increase each decade beginning with 1910. Uniformity in rate of increase is apparent, being represented by a straight line.

FIGURE 81.—Reruling the Grid of Figure 80 to Represent Round Numbers

By erasing the equidistant lines represented by 100, 110, 121, 133, etc., the grid takes on the appearance of the regular logarithmic or ratio grid. The straight line, "A," increases at a uniform ratio in equal periods of time and line, "B," which curves slightly downward, increases by equal differences.

In this connection, it should be noted that the logarithmic scale lends itself chiefly to the curve-graph. It reflects only the ratios between the numbers plotted: its position on the grid, therefore, has no particular significance. Since the bar-graph implies a significance in the length of the bar it should not be generally used in this form of graph.

3. The logarithmic scale has no zero point. This is apparent from its nature. Its chief function is to reflect the ratios or rates of change between numbers and no such computations can be made when one of the numbers is zero. Whenever zero enters into the process of multiplication the result is zero; whenever it enters into division the result is either zero or infinity. It is clear therefore that zero cannot enter into the graduations, however finely they are made or however far extended.

   It will be noted also that the process of constructing the logarithmic scale eliminates the zero point. The spacings on the logarithmic scale represent the logarithms of the numbers 1 to 10 plotted on an arithmetic scale which begins with zero. To facilitate the use of the scale, the corresponding numbers are substituted for the logarithms. Following this procedure, 1, whose logarithm is zero, is substituted for zero.

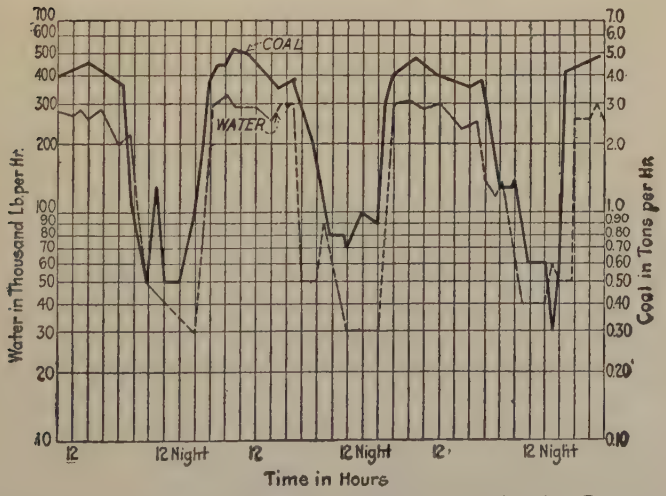
4. It follows from this, that negative numbers cannot be represented logarithmically on the same grid on which positive numbers are shown. We cannot pass from positive to negative numbers or vice versa without registering a zero point. When the zero point in a series of data is arbitrary rather than actual it may be eliminated and the data plotted on the logarithmic field as discussed on pages 407-410 and 432 of Karsten's "Charts and Graphs." The suggestion offered on page 60 of Haskell's "Graphic Charts in Business" is also interesting in this connection.

5. The scale designations may be multiplied or divided by any constant number without affecting the character of the graph based on them. This is usually necessary in order to meet the requirements of the particular series of data. In practical work it is more convenient to increase or decrease the scale by a power of 10. Thus, the scale may extend from 0.1 to 1, 1 to 10, 10 to 100, 100 to 1000, and so forth.

6. The ordinary logarithmic grid is made up of sets of rulings graduated in accordance with the series of logarithmic values of 1 to 10. Each set is called a "deck" or "cycle." While

the rulings repeat regularly from one cycle to another, the scale designations change. The designations generally increase regularly by some multiple of 10 as is shown in Figure 83, page 335, in which the first cycle runs from 10 to 100, and the second from 100 to 1000. It is apparent that the cycles will be of the same size—*i.e.*, the total distance from 10 to 100, 100 to 1000, and so forth, are equal. This is essential to the use of two or more cycles.

7. Contrary to the case of the arithmetic grid, two or more scales may be used on a single cycle of the logarithmic grid and although the points plotted may vary materially in value, the graphs are nevertheless comparable. This follows from the fact that equal vertical distances on the logarithmic scale represent equal ratios regardless of their position on the scale. This permits the plotting on a single cycle of different series of data which vary greatly in value. Figure 82 is illustrative. As pointed out on page 57 of Haskell's "Graphic Charts in Business," it also permits us to confine to a single cycle the graph of a single series of data whose range extends far beyond the range of the single cycle. We simply carry the curve to the top of the cycle with an appropriate scale, and then beginning again at a point on the



(From article by M. C. W. Tomlinson, *Management Engineering*, Dec. 1921)

FIGURE 82.—Water and Coal Consumption of a Large Heating Station
Use of two scales on logarithmic grid.

bottom of the cycle which is in line with the point where the curve ran off the grid at the top we continue the plotting with a new scale selected to suit the data. This process is repeated until the curve is finally completed.

How to Construct the Logarithmic Graph.—The steps outlined on page 284 to be followed in the construction of the arithmetic curve-graph will apply generally to the logarithmic forms. There are certain detailed points however which require special attention.

In the selection of the vertical scale, we have to determine the number of cycles that will be required in the particular case. This is, "based upon the range of fluctuation in the series to be plotted. Before determining upon the number of decks to use in your chart, you must first glance through your series and note not merely its highest but also its lowest items. If both have the same number of (integral) digits, a single deck is sufficient; but if the maximum has more (integral) digits than the minimum, you need as many additional decks as there are additional digits in the maximum figure. A little thought will enable you to determine exactly the number of decks you need before you start your chart." (Karsten's Charts and Graphs, page 393.)

For example, a series ranging from 13 to 87 will require a single cycle, 13 to 187 two cycles, 13 to 12476 four cycles, and so forth. Series ranging from 6 to 10, 40 to 100, and similar series whose upper limit corresponds with the upper limit of the cycle, will require only a single cycle. In using more than one cycle in the same graph, they must be uniform in size. This is also true in the case of different graphs which are to be compared. Graphs constructed on logarithmic grids the cycles of which differ in size are not comparable. The student should demonstrate this point by plotting the same data on cycles of different width.

The relation between the horizontal and vertical scale is quite as important in the construction of the logarithmic graph as in the arithmetic—*i.e.*, the size of the cycle and the graduations on the horizontal scale must be carefully compared. The general appearance of the graph can be changed materially merely by shifting this relationship. Following the procedure of Figure 45, page 293, the student should construct graphs to demonstrate this point.

This will hold true for both the full and semi-logarithmic grid. The practical test suggested in the discussion of the arithmetic curve will also be serviceable here—*i.e.*, that the relation should be such that the smallest relative fluctuations of importance will be nicely registered.

It should be remembered at this point that curves representing widely different values may be readily brought into juxtaposition in this form of graph. This may be accomplished by means of the split-deck grid although the use of more than one scale on the same grid as is pointed out on page 331 is more effective. The student should likewise construct graphs to illustrate this point.

Finally, there is some difference between the process of plotting the points on the logarithmic graph and on the arithmetic graph. While the method of actually locating the points is not greatly different from the corresponding process in the arithmetic curve, there is some difference in interpolating. The distances between the rulings on the logarithmic scale are graduated logarithmically while in the arithmetic scale similar spaces are graduated arithmetically. It will be observed, for example, that the point 150 in Figure 75, page 327, is more than half the distance from 100 to 200. In the arithmetic grid it would be just half-way. Likewise, the point 250 is more than half the distance from 200 to 300. However, such a distinction has little practical value, since there is little difference in the logarithmic and arithmetic scale over such short ranges.

How to Read the Logarithmic Curve.—Three different kinds of facts may be read directly from the logarithmic curve —(1) the absolute numbers which are read directly from the scale as in the arithmetic curve, (2) the percentages of change or, the ratios, which are estimated by inspection or determined more accurately by means of a scale, and (3) the rates of change, based primarily on the direction or slope of the curve.

The method of determining the absolute numbers is simple. The actual percentages of increase or decrease, or, the ratios, will be a little more difficult. It is usually preferable, particularly for beginners to construct a percentage scale to facilitate readings of

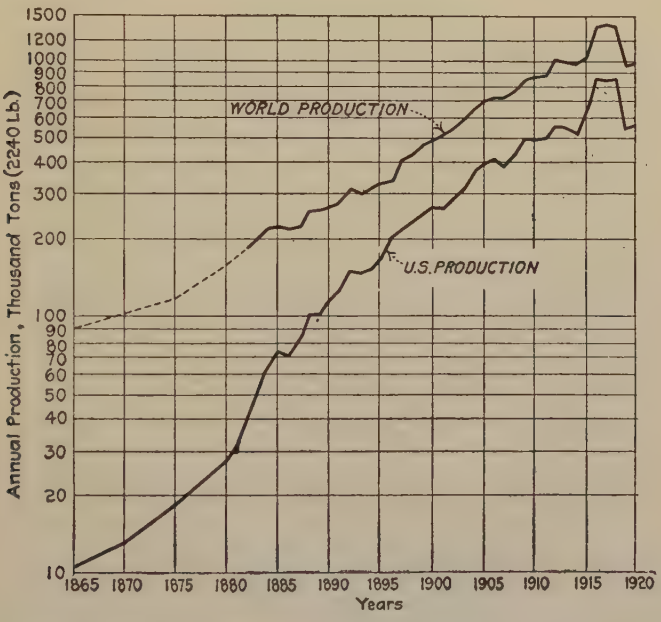
this type. As the same percentage change will be represented by different distances on different scales, a separate percentage scale will have to be constructed for each separate chart where the scales differ. Detailed discussion of this point will be found on pages 51 to 54 of Haskell's "Graphic Charts in Business."

The interpretations of the direction or slope of the logarithmic curve or portions thereof are summarized as follows on page 50 of Haskell's book:

1. If the curve is horizontal, neither ascending nor descending, the magnitudes are uniform, neither increasing nor decreasing.
2. If the curve ascends and is straight, or very nearly so, the magnitudes are increasing at a uniform, or very nearly uniform, rate.
3. If the curve ascends and bends upward, concave fashion, the magnitudes are increasing at an increasing rate.
4. If the curve ascends and bends downward, convex fashion, the magnitudes are increasing at a decreasing rate.
5. If the curve descends and is straight, or very nearly so, the magnitudes are decreasing at a uniform, or very nearly uniform, rate.
6. If the curve descends and bends upward, concave fashion, the magnitudes are decreasing at a decreasing rate.
7. If the curve descends and bends downward, convex fashion, the magnitudes are decreasing at an increasing rate.
8. If the direction of the curve in one portion is the same as that in another portion, the percentage rate of change is the same for both portions.
9. If the curve is steeper in one portion than in another portion the rate of change is more rapid in the former than in the latter.
10. If two curves on the same ratio chart are parallel to each other the same rate of change exists in both.
11. If one curve is steeper than another the rate of change in the former is more rapid than in the latter.
12. A straight line, either actual or imaginary, drawn so as to represent the general trend of a curve will show by its slope the average rate of increase or decrease of the curve; and the deviations from this average may be readily perceived.

13. A straight line, either actual or imaginary, may be drawn between any two points on the curve, between which it is desired to know the rate of change, and the slope of this line will show the rate of increase or decrease between the two points.

How the Ratio Graph is Used.—From the broad point of view, the logarithmic curve is used whenever the aim is to represent ratios or rates of change. Specifically, in practical work, it is used most extensively in the presentation of historical data. In such cases, we generally use the semi-logarithmic grid with the



(From article by J. E. Morrison, *Management Engineering*, Jan. 1922)

FIGURE 83.—The Production of Copper Since 1865

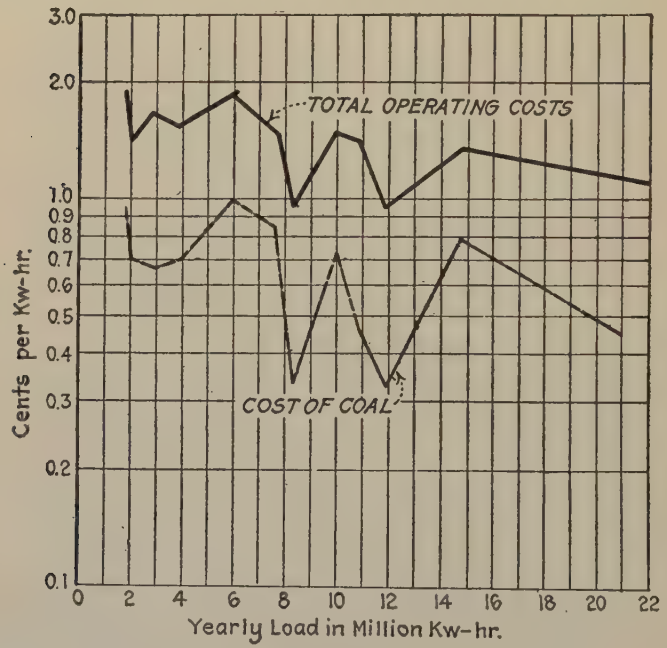
Typical logarithmic chart.

logarithmic scale on only the vertical or y-axis. For the most part the forms of paper which may be secured commercially will be quite satisfactory for this purpose.

The logarithmic curve will be used in presenting personnel, financial or operations data. Typical of the kinds of historical

facts that may well be presented logarithmically are the number of employes, labor losses, absences, wages, expenses, revenues, net profits, production output, sales and materials purchased. The advisability of presenting the data logarithmically will have to be decided in each specific case. Typical graphs of this form are represented in Figures 83 and 84.

Comparative Value of the Logarithmic or Ratio Graph.— The value of the logarithmic curve may be considered from two



(From article by M. C. W. Tomlinson, *Management Engineering*, Dec. 1921)

FIGURE 84.—Curves Showing the Influence of the Cost of Coal on the Total Operating Cost of a Large Power Plant

Typical logarithmic chart.

aspects. It may be compared in the first place with other methods of presenting ratios and rates of change and, in the second place, the comparison may be made with the arithmetic or difference curve.

Ratios and rates of change in a series of data may be represented (1) by actually making these computations and plotting them on an arithmetic scale, or (2) by devising a scale on the grid so that the same results will be obtained by plotting the absolute numbers. The latter method has certain distinct advantages. It shows the actual values, the relative values, the actual changes, and the relative changes while only a part of this information would be shown by the graph based on the computations; it avoids the necessity of selecting a base as is required in the computation of relatives; and it avoids the work of computation which in an elaborate series may be quite a burden.

In the second case, the logarithmic curve is not a substitute for the arithmetic. They are designed to serve two separate purposes and consequently should supplement each other. It is frequently advisable to present the two at the same time, the one representing the comparison of the absolute amounts and the other reflecting the ratios and rates of change. As Mr. Karsten points out on page 421 *ff.* of his "Charts and Graphs," both forms of curve may be presented on the same grid. In such cases, the arithmetic scale should usually be given greatest prominence. Methods of constructing the two curves on a single grid are discussed by Mr. Karsten in the reference just noted.

CHAPTER XXXII

MAPS—CIRCLES—AREAS—VOLUMES—PICTURES

Fundamental Facts of Area and Volume Graphs.—The principles outlined in the preceding chapter will apply fundamentally to all types of graphic devices. Their application is not so direct as in the construction of the curve but the same general methods of procedure will be apparent on careful inspection. In this chapter we shall consider the method of constructing statistical graphs in which the data are represented by map, circle, area, volume, or picture graphs. Generally speaking these types of graphs are not used extensively and should not be considered as substitutes for the curve or bar. The discussion will be comparatively brief: those who are interested further are referred to the references listed on page 351.

Fundamentally, the area and volume may be constructed in two different ways. The data may be represented by a single dimension as the side of a square or rectangle, the diameter of a circle or the edge of a cube or parallelogram. Or, the data may be represented by the area or volume, the dimensions having no direct relation to the representation of the data. If the first form is used, the method of construction is simple—the dimension of the graph is drawn to represent the data. If the second is used the construction is more difficult. Obviously, if the side of the square or diameter of the circle is doubled the area will be increased four times and if the side of the cube or the diameter of the sphere is doubled the volume will be increased nine times. Hence, if the area in the one case and the volume in the other is to be doubled it is clear that the dimensions of the squares and circles must be in the ratio of 1 to $\sqrt{2}$ or approximately 1 to 1.41 and those of the volumes, 1 to approximately 1.3.

This form of graph requires comparatively extensive computation and the result contributes little to the end that the graph is supposed to serve. It does not clarify. It is very difficult to compare graphs which involve two or more dimensions.

It is possible, of course, to have a combination of these two forms so that the dimensions represent certain facts and the area or volume another figure derived from the first. The sides of a square or rectangle for example, may represent the average money value of a sale and the number of sales for a month and the area —the product of these two—will consequently represent the total value of sales for the month. Volumes may be used in a similar manner. There is very little use for devices of this kind in practical business statistics.

The proper evaluation of these types of graphs generally will condemn them for extensive use. The single dimension graph is by far the most effective and the great bulk of graphic work in the practical business concern will be most effectively accomplished through the use of the column or bar and the curve. It is a fundamental principle that graphs are effective in reverse order of the number of dimensions involved in the representation of the data. The single dimension graph is so far superior that it should be used whenever it is possible to do so.

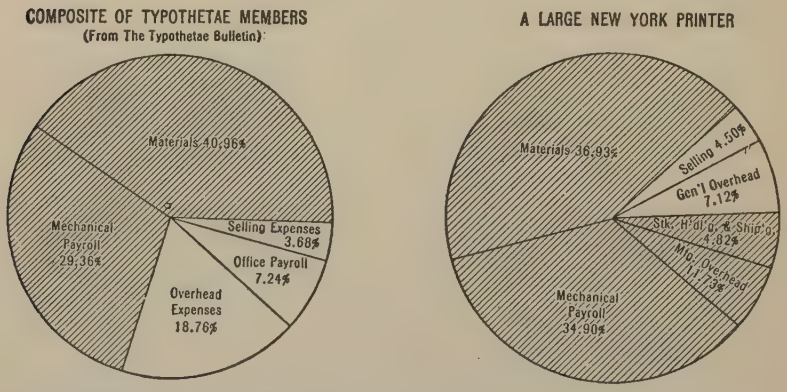
For special reasons, the cartogram and a certain form of the circle are used in connection with certain types of presentations and the picture graph is used to serve a specific purpose. We shall consider in some detail the method of constructing these three forms of graphic devices.

The Circle or Pie Graph.—In this form of graph, a scale is constructed on the circumference of a circle and the data to be presented are plotted on this scale. A point at the top of the circle is arbitrarily taken as zero or the starting point and the scale laid off from this point, clock-wise around the circle. The scale will always take the form of dividing the total circumference up into equal parts. To facilitate the recording of the data, the circle is usually divided into 100 equal parts or spaces representing 1%.

It is apparent therefore that the use of this graph is confined to the presentation of data reflecting the component parts of a whole. In plotting the points the data to be presented are preferably expressed in percentages of the total. The location of the points consists merely in assigning to any figure the number of

spaces on the scale represented by its percentage of the total number. We start with zero and count off so many spaces clockwise to represent one—generally the largest—of the component parts. Beginning then at this point we again count off enough spaces to represent a second component part—generally the second largest—and so forth.

At each of the boundary lines between component parts a radius of the circle is drawn or the point on the circumference is connected with the center of the circle by a straight line. We then have the circle divided into sectors corresponding to the component parts of the total figure which was to be represented. Finally, these



(From article by Arthur R. Burnet, *Administration*, Dec. 1921)

FIGURE 85.—Charts Showing the Analysis of the Total Cost of a Printed Product, Expressed in Units of $100

Typical circle chart.

sectors may be distinguished by different colors or shadings or they may be left open. The graph designations may be entered in the body of the sector or on the circumference just outside. If the labeling is so extensive that it is difficult to enter, a separate key will be preferable. It is clear that the data may be represented by three different parts of this graph—the arc of the circle, the subtended angle, or the area of the sector. While theoretically we assume the first of these to be preferable, as a matter of practice we are misled into comparing areas.

The circle-graph is not to be recommended for general use. In fact we may say that its use should be confined to a popular

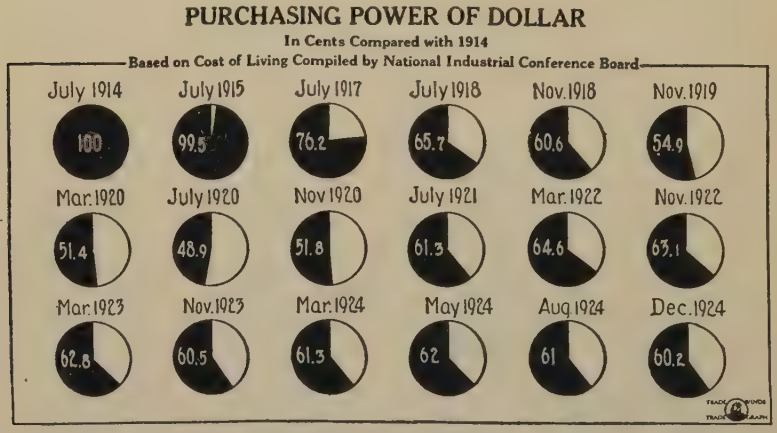
or semi-popular purpose where fine comparisons between the different parts are not required. It is very difficult to make ready comparisons between arcs, angles or areas, but where the general effect is the principal aim it will serve. A favorite use is in the representation of the component parts of the dollar either in revenues or expenses. It is also used in showing the component parts of the cost of living, the component elements which enter into the manufacture of a product, and so forth. It is very effective in some cases of this kind: it at once suggests a whole divided into parts. Its use as a progress graph is familiar. Adaptations of the simple form of the circle-graph have produced a variety of charts. Figures 85 to 87 show the more standard forms.

(Courtesy of New York *Tribune* and Chicago *Tribune*)

FIGURE 86.—Distribution of a Dollar Spent for Men's Clothing
Typical circle chart.

The Statistical Map.—The statistical map or cartogram is nothing more than an ordinary map with statistical data graphically presented and associated with the territory on the map to which they relate. The distinctive feature of the statistical map is its concrete representation of position. The spatial significance of the data is portrayed by the map and the mind is relieved of the burden of visualizing the territory to which the data are related. The real problem becomes one of representing the data on the map so that different territories may be readily distinguished or compared. The data are at hand for certain geographical units such

as sections, townships, counties, groups of counties, states, and groups of states. They will usually be reduced to a per unit basis such as the population per square mile, sales per population, commodity consumption per capita, taxes per capita, and so forth. When this has been done the data are arrayed in ascending order and the highest and lowest figure noted. On the basis of this array the data are grouped into broader classes, the number depending on the character of the data and the purpose to be served in the particular case. It should be emphasized that the classes or groups should be of the same width or size.



This chart, prepared by the Union Trust Company of Cleveland, indicates the changes in the purchasing power of the dollar from 1914 to 1925.

(From *Forbes* by permission)

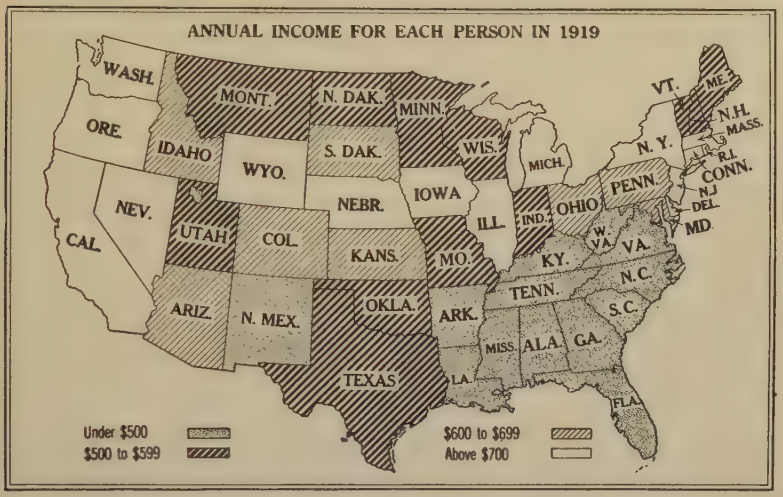
FIGURE 87.—Typical Circle Chart

There are different ways of representing these data on the map depending on the type of graphic mechanism selected to bring out the distinctions. Chief among these devices are colors, cross-hatching or shading, dots, and pins and ribbons, giving us what are called:

1. Color maps
2. Cross-hatch or shaded maps
3. Dot maps
4. Pin and ribbon maps

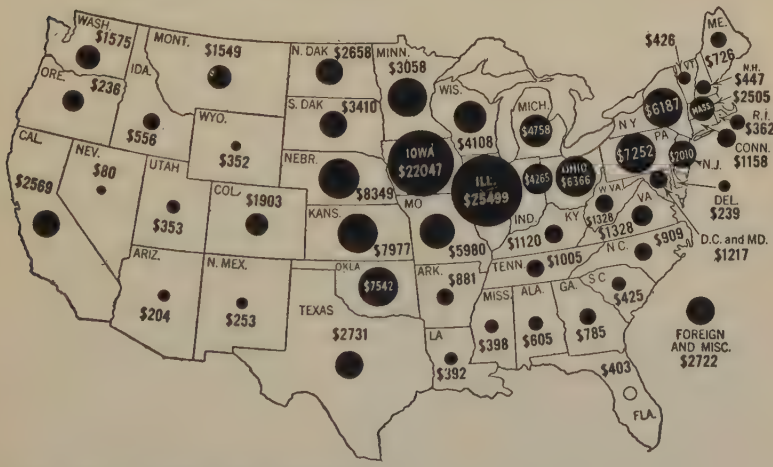
In addition to these the graph may be put up in a more popular

form by inserting bars, circles or picture-graphs in the territorial units as discussed on page 349.



(From Cleveland Trust Co. *Business Bulletin* by permission)

FIGURE 88.—Cartogram with System of Cross-hatching Representing Data

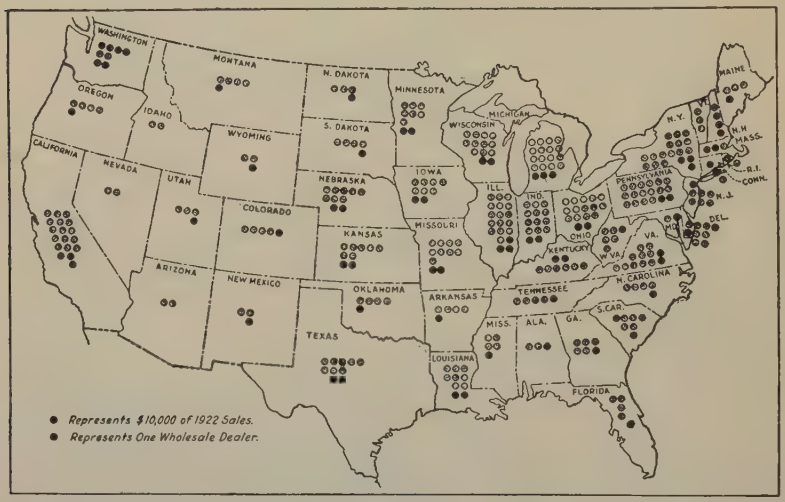


(From Chicago *Tribune Book of Facts*)

FIGURE 89.—Cartogram with Size of Dot Representing Data

Color and Cross-Hatch Maps.—The primary step in the construction of the statistical map is to assign to each of the groups of data a distinct graphic scheme. If colors are used, different colors or different intensities of the same color will be assigned to each group of data: or if cross-hatching or shading is used, a distinct form of cross-hatching or shading will be so assigned. Figure 88 above will serve as an illustration.

If dots are used as the medium we may select one of three different schemes. First, we may allow the size of the dot to reflect the value of the data. In this case we will determine a scale according to which a certain size of dot will represent a certain value and then on the basis of this scale prepare dots to represent



(From article by H. N. Stronck, *Management and Administration*, June 1924)

FIGURE 90.—A Typical Geographical Chart Showing the Distribution of 1922 Sales in Dollars and of Wholesale Dealers in the United States

Data represented by frequency of dots.

the several data, placing the dots so determined in their respective geographical units. Figure 89 above will illustrate this type. In the construction of the dots the basis of representation may be their dimensions or their area. In either case they are of little value in comparisons involving a fine degree of accuracy.

Second, we may select one size of dot and depend on its fre-

quency on the map to reflect the value of the data. In this case the scale to assign to the dot must receive careful consideration. If it is too small we may find the dots too congested in the territories which have the highest figures and if it is too large those in the lowest territory may not appear at all. The problem is somewhat similar to that of selecting the scale for the curve-graph. Figure 90 above is illustrative.

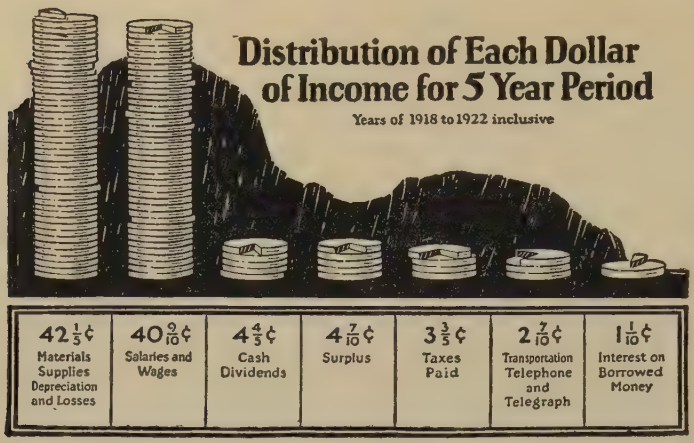
Finally we may use one size of dot and use a system of shading or cross-hatching to reflect the value of the data. Here the size of the dot must be large enough to make the distinctive marking apparent. Its use is advisable in only rare cases.

While this type of graph may be used in the presentation of personnel, financial, or operations facts generally, its most extensive use in practical business statistics is in the presentation of marketing facts.

Pin and Ribbon Maps.—Pins may be used to represent the data on the map instead of the colors, shadings and dots. These are particularly serviceable where changing conditions are to be portrayed, such as the actual conditions of sales on each day. The tacks or pins come in a great variety of colors, designs, and sizes and consequently a great variety of data can be represented by them. The service from them is not different from the use of the other devices noted above except that they are much more flexible. In Figure 90 above, for example, pins may be used instead of the dots. Much may be accomplished also by the use of the different colored ribbons with the pins or tacks.

The most satisfactory method of procedure in considering the use of these devices in practical work is to communicate directly with their distributors such as the Educational Exhibition Company, Providence, R. I.; Codex Book Company, New York City; Hammond Map Company, New York City; Rand-McNally Company, New York City; and others.

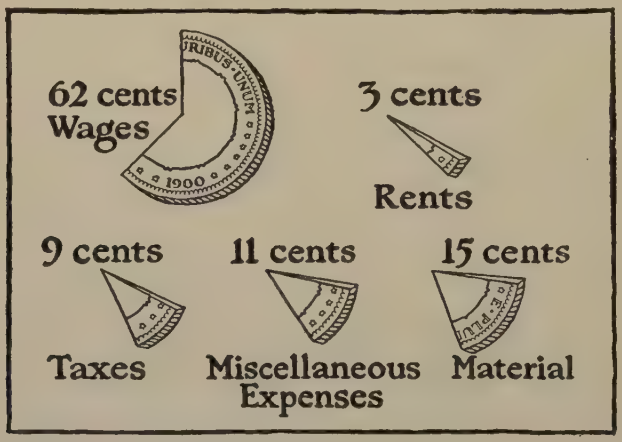
Picture-Graphs.—In the picture-graph the data are represented by pictures of the things for which they are the numerical symbols. Statistics on coal are represented by piles of coal, building data by houses, bread by loaves, meat by cattle or hogs, and so forth. The charts given, Figures 91 to 96, are typical.


Distribution of Each Dollar of Income for 5 Year Period
Years of 1918 to 1922 inclusive
42 1/3¢ Materials Supplies Depreciation and Losses
40 9/10¢ Salaries and Wages
4 4/5¢ Cash Dividends
4 7/10¢ Surplus
3 3/5¢ Taxes Paid
2 7/10¢ Transportation Telephone and Telegraph
1 1/10¢ Interest on Borrowed Money

(Courtesy General Electric Co. and *Saturday Evening Post*)

FIGURE 91

The distinctive feature about this type of graph is its pictorial appeal. The essential elements of its construction are not dif-

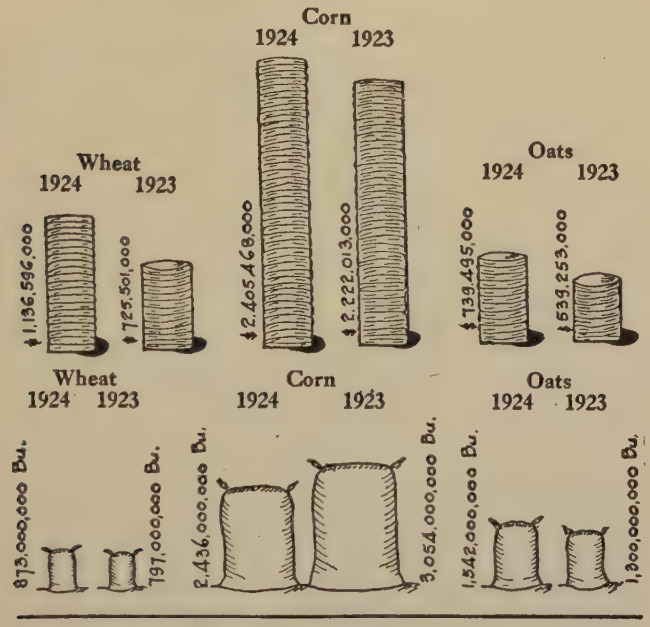


(Courtesy of the *Telephone Almanac*)

FIGURE 92.—Typical Disposition of a Dollar Expended for Telephone Operating Expenses

ferent from those of other types: it is evident that a single dimension, an area, or volume may be the basis of constructing the

picture-graph. Generally, the single dimension is used and it is by far preferable. The heights of men or houses and the lengths of cars or ships should represent the data and the rest of the



(From *The Industrial Digest* by permission)

FIGURE 93.—Production of Wheat, Corn and Oats—1923 and 1924 Compared

picture in each case be made to conform to this single dimension. The usual method of developing the picture-graph is to represent

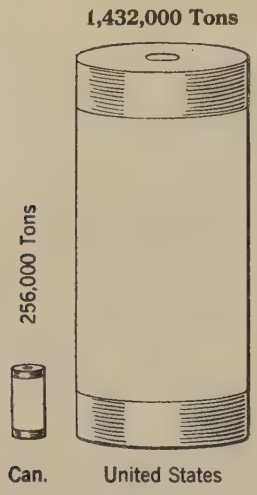
(From *The Industrial Digest* by permission)

FIGURE 94.—Production of Motor Cars First 11 Months of the Years Shown

each figure by a single picture, the difference in the sizes of the pictures reflecting the difference in the magnitudes of the data.

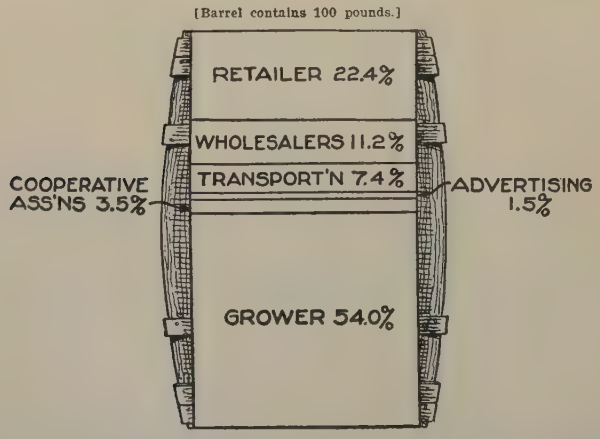
All sorts of picture-graphs have been developed for all sorts of purposes in personnel, financial, and operations facts and the

field of their application is rapidly expanding. It should be emphasized that their contribution is strictly in the popular field;



(From *The Industrial Digest* by permission)

FIGURE 95.—Newsprint Production in United States and Canada

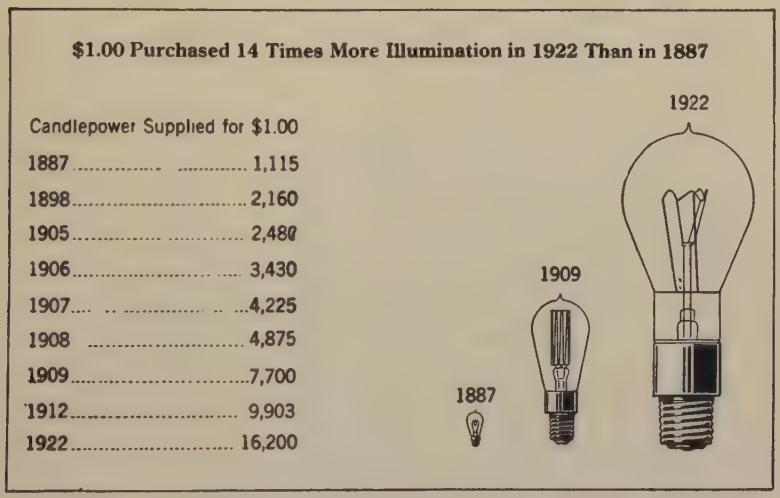


(From *Report of Joint Commission of Agricultural Inquiry* by permission)

FIGURE 96.—Analysis of Consumers' Price of a Barrel of Cranberries—1920

they have little to offer in cases where a fine degree of accuracy is sought. As illustrative of the different types of devices that have been worked out, Figures 91 to 96 are presented.

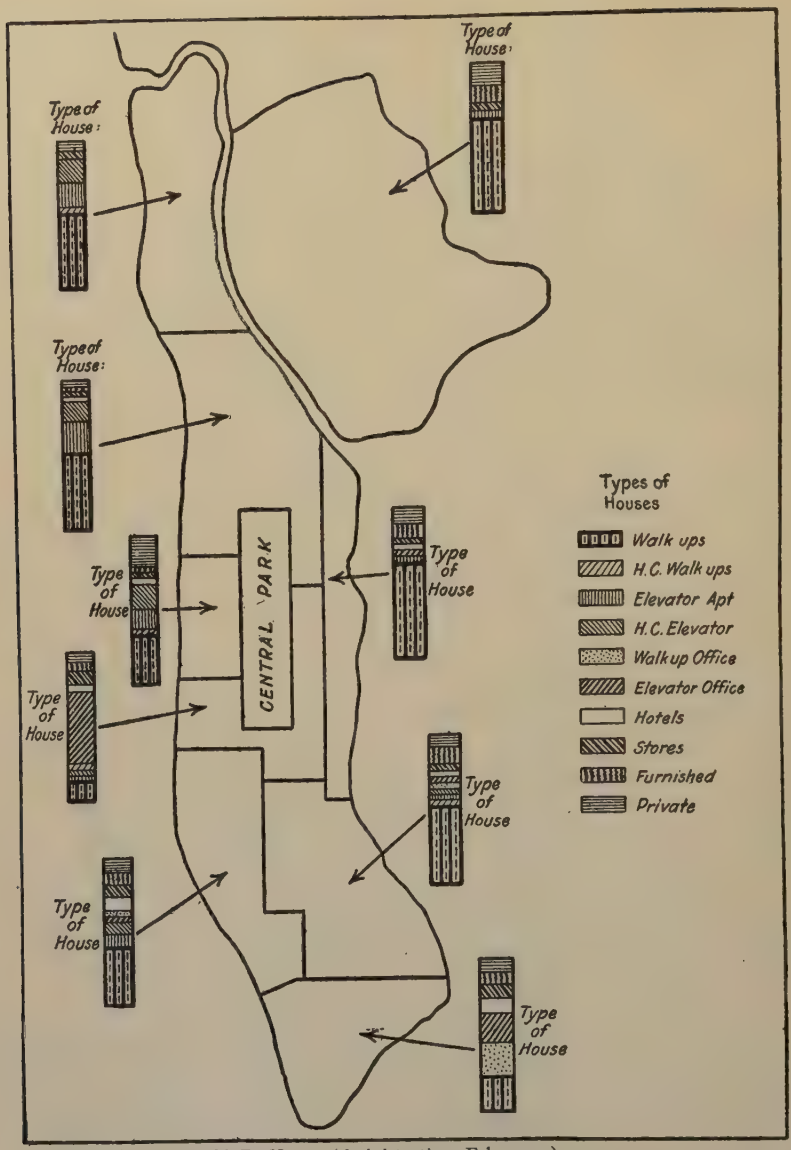
Combination Graphs.—A combination of two or more of the forms outlined above is sometimes made to effect a graphic presentation. The picture-graph particularly is frequently combined with other types, especially the bar, the curve, the map, and the circle. Figures 96 and 97 are typical. A variation is to make the pictures the same size and arrange them in rows, the number in each row being determined by the amount of the data to be represented. Since the lengths of the rows become the basis of



(Courtesy of the New York *Sun* and Harris, Forbes & Co.)

FIGURE 97

comparison, this arrangement overcomes to a certain extent the disadvantages of the picture-graph. Similarly the curve graph may be constructed by arranging a row of pictures along the line of the curve. Typical illustrations of these types may be found on pages 128 and 219 of Karsten's "Charts and Graphs." Other forms may likewise be combined as follows:



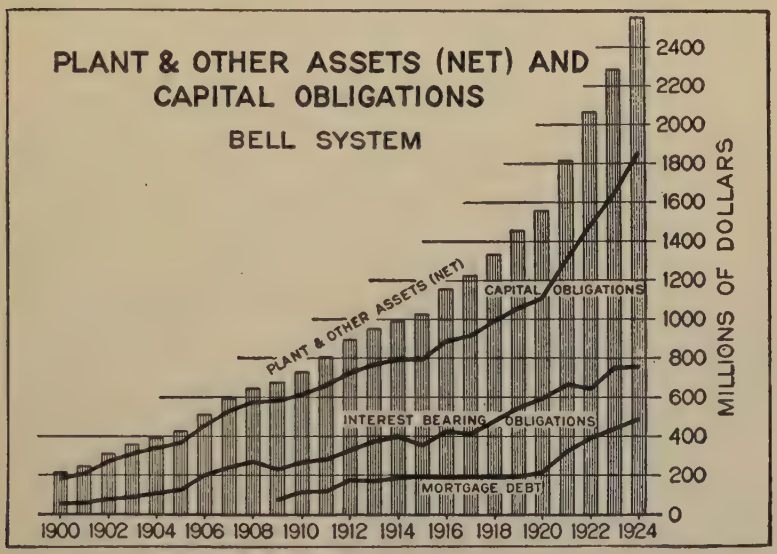
(From article by Harold B. Ness, *Administration*, Feb. 1923)

FIGURE 98.—Map of New York Divided into Delivery Zones, with the Proportion of the Various Types of Houses in Each

Combination graph—map and bars.

1. Bar and map, curve and map, or circle and map—Construct the bar, curve, or circle graph at appropriate points on the map.
2. Bar and circle—Construct the bar graph in the sectors of the circle.
3. Bar and curve—Construct the bar graph from the horizontal base and the curve graph running across the grid.

Figures 98 and 99 are typical of these forms.



(From 1924 Annual Report of American Telephone and Telegraph Co.)

FIGURE 99.—Curve and Column Graphs in Combination
Combination graph—curves and columns.

REFERENCES—CHAPTERS XXVII TO XXXII (GRAPHICS)

Brinton, W. C. Graphic Methods for Presenting Facts, 1914.
Chaney, L. W. Comparison of "Arithmetic" and "Ratio" charts—Monthly Labor Review—March, 1919, pp. 20-34.
Fisher, Irving. The "Ratio" Chart for Plotting Statistics. Journal of the American Statistical Association, June 1917, pp. 577-601.
Haskell, A. C. Graphic Charts in Business, 1924.
Jerome, Harry. Statistical Method, 1924, Chs. IV, V, and VI.
Jones, D. C. A first Course in Statistics, 1921, Chs. VIII and IX.
Karsten, K. G. Charts and Graphs, 1923.

Knoeppel, C. E. Graphic Production Control, 1920.
Manual of Charting. Business School Series, 1923.
Marshall, W. C. Graphical Methods, 1921.
Savage, W. A. Graphic Analysis for Executives, 1924.
Secrist, Horace. An Introduction to Statistical Methods, 1917, Chs. VI and VII.
—— Statistics in Business, 1920, Ch. V.
Smith, W. H. Graphic Statistics in Management, 1924.
Statistical Atlas. United States Census Bureau.
West, C. J. Introduction to Mathematical Statistics, 1917, Chs. I, II, III and VI.

Part IV—Analysis and Interpretation of Statistical Data

CHAPTER XXXIII

ANALYSIS—GENERAL

Analysis Essential.—Thus far in these discussions we have been considering the collection of the facts required in a program of statistics in a business organization, the systematic recording of these facts, and the methods of presenting them. This brings us to the subject of the analysis of business facts which involves the determination of the significance of the facts in the light of the important problems of the business. The detailed facts which have been provided must be analyzed to serve the needs of the management. This is the real purpose in collecting and systematically recording the detailed information suggested in Chapters III to VIII and the story is only half-told when the facts have been made available. Statistical data and statistical methods are but means to an end and that end in business statistics is to aid the management in solving the problems with which they are confronted in running the business. The facts must be analyzed if they are to render this service.

As emphasized on page 231, the treatment will take the form of explaining separately each method of analysis and of indicating its use in practical business statistics.

The Meaning of Analysis.—Analysis signifies a breaking up into component parts. Concretely, for our purposes, it may be defined as the processes of preparing statistical data for interpretation by means of (1) classifying them, (2) associating them, and (3) developing them into various forms through computations such as percentages, ratios, averages, index numbers and so forth. Analysis breaks groups up into component parts, sum-

mates them, combines them, changes their forms, brings them together in association or relationship, and otherwise treats them so as to reveal their significance.

The Aim of Analysis.—The final aim in the development of statistical studies in a business organization is the interpretation or evaluation of statistics in the light of the problems of the business. This point is amplified later, in the discussion of interpretation. Now, analysis is designed primarily for the purpose of facilitating interpretation. The fundamental aim in analysis is to put the facts in a form which will make it possible to interpret them with a minimum amount of time and energy. Analysis is merely a means to an end and the end is the proper interpretation of the facts from the standpoint of the management of the business. It simply sets the data or facts up in such way that their meaning may be most readily and beneficially determined. Interpretation which may be done by the statistician or by the executives to whom the facts are presented, follows analysis. It is discussed in Chapter XLIV beginning on page 484. The importance of pointing out the conclusions resulting from analysis is also emphasized in that chapter.

Processes of Analysis.—In accomplishing analysis three broad processes are followed—(1) classification in which the facts having like characteristics are grouped together, (2) association in which different groups of facts are related or associated in orderly and significant manner, and (3) computation, in which the data are summated or in which percentages, ratios, averages, indexes, and so forth, are computed. These processes may be used independently or they may be used concurrently.

Analysis by Classification.—Classification consists primarily of arranging like facts in groups and is the first and fundamental step in the process of analysis. It is used from the very beginning through to the end of the development of a program of statistics. Even in the preparation for the collection of data classification plays an important part and in building up systematic basic records as well as in the final analysis of the facts it is one of the most effective tools.

The fact, however, that classification enters into practically every phase of statistical development should not becloud its true nature. Wherever it is used, it is an analytical device just as truly as averages, index numbers, and other computations developed in later chapters and its aim is the same as that of any other analytical process—to reveal important conclusions which may be established from the data. Business facts are thrown into groups because they are generally alike in their bearing on the problems of the business. The importance of time and space classifications, for example, lies in the fact that comparison of time periods or special units is an effective means of revealing important conclusions from the data.

Primary Bases of Classification.—By way of suggesting a comprehensive picture of the primary classes of data, for practical purposes it may be serviceable to classify statistical data on two broad bases—(1) characteristics or attributes of the statistical unit, and (2) elements of the situation in which the observation is made. In more detail these groups are as follows:

I. Characteristics or attributes of the statistical unit
   1. Qualitative—In which the basis of classification is the quality or character of the attribute.
   2. Quantitative—In which the basis of classification is the degree or measurement of the attribute.

II. Elements of the situation
   1. Time
   2. Space
   3. Other

These bases of classification are amplified below.

It should be noted at this point that classification is generally made on the basis of a single element or attribute. Two facts may be put in the same class when the classification is on one basis and in entirely different classes when it is on another basis. For instance, two employes may be placed in the same class when occupation is the basis of classification and in widely different classes when age or rate of pay is the basis.

Qualitative Classification.—The basis of this type of classification is a qualitative attribute or characteristic of the statistical

unit. In general outline, qualitative classification consists of the establishment first of broad classes and then of breaking these classes up into such sub-classes and sub-sub-classes as may seem advisable. Thus in assembling internal data of the business organization, the data are first segregated into personnel, financial, and operations groups. The personnel data are then sub-grouped into *the working force, the wages,* and *the employment conditions;* financial data into *the acquisition of capital, the use of capital,* and *the return on capital;* and operations data into *purchasing facts, producing facts,* and *marketing facts.* Further classification is outlined in Chapters III, V, and VII where the qualitative grouping is carried out into great detail—the employe, for example, being classified as to the qualities of sex, nationality, occupation, education, health, and so forth. The specific classes in any given case will be determined by the properties of the statistical unit and the purpose to be served in that case.

Quantitative Classification.—Quantitative classification is based on the measurement, numerically expressed, of the statistical unit. Whether the unit is dollars, pounds, inches, products, sales, deliveries, and so forth, depends on the particular situation. Quantitative classification presupposes qualitative classification. It involves for the most part the grouping of qualitative measurements of the statistical unit into measurement classes, just as qualitative classification consists of so grouping units which have common qualitative properties. Qualitative classification will throw together items of like kind—*e.g.,* wages with wages, prices with prices, and sales with sales. Further homogeneity will be established by sub-grouping these data into different kinds. But even this is not sufficient. We want to know not only that a fact is a price of a certain kind but we also want to know the value of the price: that is, we want a quantitative measurement.

Two principal considerations are involved in quantitative classification. First, the fineness of the measurement observed in collecting the data or making the observations, and, second, the grouping of these measurements into broader classes for analytical purposes. For example, in collecting wage facts, the data may

represent weekly earnings to the nearest dollar, half-dollar, or cent, and the data may then be grouped into dollar, dollar-and-a-half, two-dollars-and-a-half, or five-dollar classes.

It is obvious that the first of these points will have to be disposed of before the data are collected. We must determine in what unit of measurement the data are to be expressed and to how fine a degree of accuracy the measurements are to be taken. Quantitative classification, however, has to do primarily with the second point—*i.e.*, the organization of the detailed measurements into broader classes. In the regular program of statistics, the most detailed measurements required must be maintained in the basic records.

The classification of quantitative data is given special consideration in the discussion of *frequency analysis* in Chapter XXXV, beginning on page 377.

Elements of the Situation.—The significance of statistical data will often vary materially with the situation in which the observations were made that produced the data. Consequently, the elements of the situation form an important basis of classification. Obviously, it would be impossible to make a comprehensive list of these elements. Those which are perhaps most generally present and which more frequently form the basis of classification in practical work are *time* and *place*. An observation is made at a specific place and at a certain time and it is these elements of the situation which are generally of primary interest. However, there are others of importance as pointed out below.

Time Classification.—The time element or chronology is the controlling factor in this type of classification. Data are arranged in classes in accordance with certain periods of time—years, months, or days—as illustrated by Table 1, page 366.

This historical type of classification is extensively used in business statistics. It is developed to indicate the trend of earnings and expenses, the payroll, number of employes, and various other facts, internal and external, which are important in the direction of the business. This type of classification requires

certain special analytical treatment as developed in Chapter XLIII, beginning on page 471.

Space Classification.—The basic element in this classification is the geographical factor. Data are arranged by geographical territory, indicated by whatever units are appropriate in the specific case. We may classify population data by counties or states, sales data by sales districts, earnings and expenses by territorial divisions, and so forth.

This method of classification is very familiar and is used freely in the field of practical business statistics. No particular phases of it require special treatment. Generally, the only problems involved are those arising in connection with the determination of the territorial boundaries and the decision as to what data belong in each territory.

With internal data the territorial classification is fixed by the company organization, the problems under consideration usually not involving a different territorial or geographical classification. With external data, however, the selection of geographic units is left largely to the judgment of the statistician. It is not always an easy matter to decide this question. There is danger of innocently coloring the facts through improperly selecting the territorial sections and there is likewise the possibility of "statistical gerrymandering," that is to say, so grouping favorable and unfavorable territories as to make the results satisfactory even though the data do not represent the actual conditions. For example, in making a study of the foreign born population in the United States we should get entirely different results, by confining the inquiry to states which are largely rural, from those obtained by making the same sort of study in states which are largely urban. In studying the labor market with relation to a business concern, the geographical territory from which the concern attracts labor actually or potentially must be carefully determined. Table 4, Chapter XXV, is typical of this type of classification.

Other Elements of the Situation.—Other elements of importance in the situation will vary widely with the difference in the detailed character of the situations in which observations are

made. Frequently weather and atmospheric conditions are important as, for example, in facts relative to local deliveries where dry, slippery, or snow-covered streets may be significant; in facts relative to the use of trolley cars, subways, taxis and local telephones; in facts relative to telephonic and radio communication; in facts relative to the number of customers in general and the calls for special classes of goods in the retailing business; and so forth. Other factors of equal importance will be present in different situations and frequently furnish a desirable basis of classification.

CLASSIFICATION: SUMMARY.—The extent to which classification should be carried will be governed by the facts of the particular situation. It may vary from the simple two-fold grouping to a manifold grouping with a multitude of attributes or elements involved. The outline in Chapters III to VIII together with the accompanying discussion will suggest the possibilities. The following taken from Hatfield's "Modern Accounting" is also suggestive:

> Thus the differentiation is continued, real property separated from personal, merchandise distinguished from plant, and still further subdivided into different classes, as for instance, dry goods, and groceries, each of these divided into groups such as: silk goods, linens, notions; flour, tea and coffee, sugar, and thus indefinitely according to the nature of the business, the taste of the proprietor and the particular purpose for which the accounts are kept. No rule can be laid down as to the degree to which ramifications may be made, as each additional subdivision gives to the proprietor additional information valuable in the conduct of the business. The only limit is that at some point, varying with the particular business, the expense of getting the additional information, requiring as it does more books and more clerks, is greater than the value of the information thus gained. Somewhat before this point the process should stop.

In practical work, however, only those classes should be selected which have a practical significance. The cautions emphasized in the discussion of the collection of statistical data should be doubly emphasized here. The classes selected should conform to the natural characteristics of the data in so far as

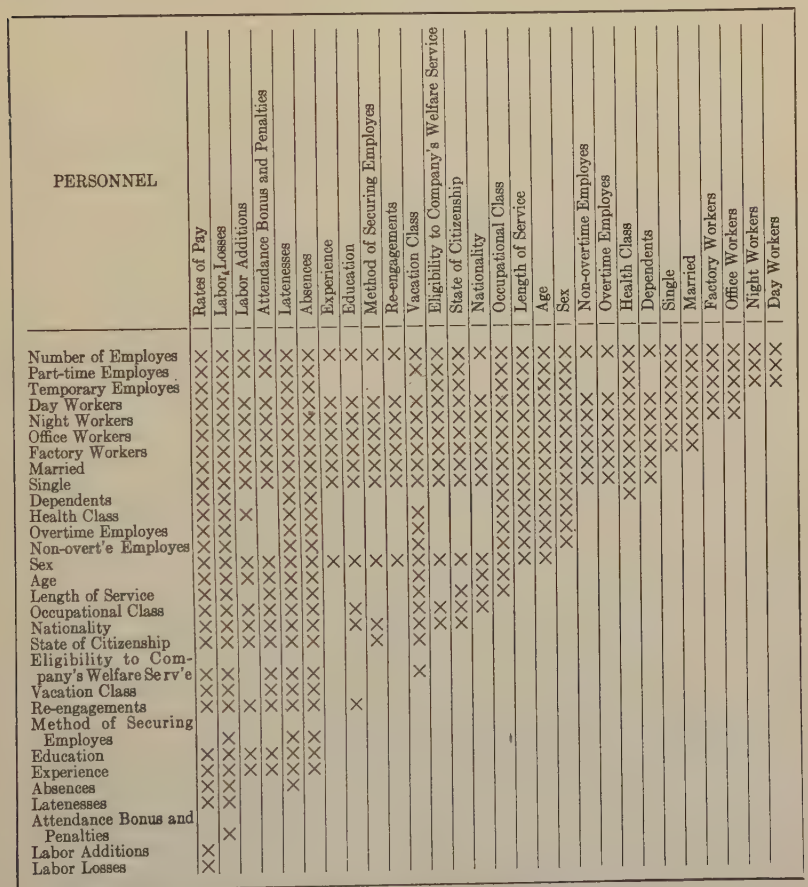
possible. In many cases, arbitrary classes are found to be set up and in such cases extreme care should be taken lest the arbitrary decisions distort the true meaning of the data. Any serious difficulties involved in the classification should be carefully noted and associated with the classified data.

In detailed classifications, a miscellaneous group will frequently be provided to serve as a sort of "catch-all" for scattering and otherwise relatively unimportant items. This class may be called "unknown," "not reported," "not specified," "others," "above—," "below—," or "miscellaneous." While care must be observed in the use of this class, it is generally very serviceable. For the most part, however, the items contained in it should be reduced to a minimum consistent with effectiveness and economy in presentation. If they form a considerable part of the total number of items, we should note fully just what is contained in the class. As a matter of fact, it is usually well to indicate why such class was developed and note in a general way what it contains. Furthermore, if the items contained in the class have any considerable or special bearing on the general picture, this should be carefully noted.

Analysis by Association.—This type of analysis consists merely of a special form of classification which because of its importance is given special treatment. By association is meant the bringing of two series of data having different units together in systematic manner so as to produce a cross classification. A number of the facts listed in Chapters III to VIII take on an added significance when associated with other facts. For example, it is significant to note the number of employes in each wage class, but such figures are more important when associated with age, sex, occupational, and period of service groups. Figure 21, page 228, represents a typical association and Figure 100 below suggests the extent of association in the personnel field. Similar charts should be developed for financial and operations facts.

Analysis by Computation.—The computation of statistical data into various forms is one of the most effective and the most extensively used methods of analysis in all fields of statistical

data—personnel, financial, and operations. While arithmetic processes generally are used freely, there are certain forms of com-



NOTE—A check mark (X) indicates that the corresponding items in the vertical and horizontal lists may well be associated. For example, the check mark (X) opposite "Length of Service" in the vertical column and "Rates of Pay" in the horizontal list indicates that the wage groups of employes should be analyzed to reflect the length of service.

FIGURE 100

putations whose significance and frequent use give them a special importance in statistical work. Chief among these are summations

and cumulations, ratios, percentages, and rates of different forms, averages, index numbers, and the more technical computations such as the average deviation, the standard deviation, and the coefficients of dispersion, skewness, and correlation.

Averages are discussed at some length in Chapters XXXVI to XXXIX, beginning on page 402; index numbers in Chapter XLI, beginning on page 449; the average deviation, standard deviation, and the coefficients of dispersion and skewness in Chapter XL, beginning on page 436; and the coefficient of correlation in Chapter XLII, beginning on page 464. The other forms of computations listed will be developed at this point.

Summations.—The fundamental character of this process is quite familiar; it is nothing more than simple addition. Likewise, the meaning of sub-totals is clearly understood as shown in Table 1, page 248. There are, however, certain special applications of this process in statistical work which are worthy of note:

1. *The Weighted Total*—In this process the items in the column are multiplied severally by certain figures called weights and the products then totaled. This process consists of nothing more than the use of multiplication as a short method of addition. We simply multiply the items by their respective weights instead of listing them as many times as there are units in the weights. Table 3, page 408, shows the totaling of the data of Table 2 by straight addition. It is simply an economy of arithmetic process and will be used, whenever applicable, in all kinds of data.
2. *The Checking Total*—This refers to the single total which appears on certain forms of statistical and accounting statements and which is obtained by summating the items in a vertical column or those in its corresponding horizontal column. 57247 in Table 1, page 248 is the sum of the items listed in column "d" and also of those listed in row 17. It is an effective means of checking the accuracy of data in certain forms of tabular presentation.
3. *The Cumulative Total*—This is the total determined by adding each successive item to those preceding. The process may be carried out by beginning at either end of the column of figures. Thus in Table 1 on page 366 column "e" shows the cumulative totals from the top downward or in ascending

order. This process is used extensively in frequency classification of all kinds and in such data as the expenses and revenues for the period of a year.

4. *The Moving Total*—The moving total represents the sum of a definite number of items in the column. The actual items included, however, are progressively changed by dropping an item from one end of the group and including at the other end the next item in the column. In Table 1, page 366, column "g" contains moving totals of a group of 12 figures. The first total is the sum of the items for the months February, 1922, to January, 1923, inclusive. To get the second total we drop the month of February, 1922, and include the month of February, 1923, the sum of the months March, 1922, to February, 1923, inclusive, giving the second entry. And so the process is continued. The total may be entered opposite the middle item in the group summated, the last item in the group or the first item in the group, to suit the requirements of the particular case. The important point is to keep in mind the time periods represented by the total. The moving total is used especially to smooth an historical series and to indicate the general trend.

Ratios of Various Forms.—This form of computation is but an expression of the numerical relationship between two numbers. Different terms are used generally to designate different ways of expressing this relationship. Chief among those used in elementary business statistics are the following:

1. *Ratio*—This is the most comprehensive term. In strictness, it may be applied to any of the different computations expressing the numerical relationship between numbers, the other terms being considered as designating certain types of ratios. Specifically, however, "ratio" is commonly used to represent a definite computation—*viz.*, the simple division of one number by another. Thus, the ratio of 1 to 2 is ½, and of 75 to 100, 75/100 or ¾.

When applied to business facts, it means simply that one fact is divided by the other. Thus if current assets are $7,500 and current liabilities $2,500, the working capital ratio may be expressed by taking $7,500 as the numerator and $2,500 as the denominator of a fraction, or simply by dividing $7,500 by $2,500, which gives a ratio of 3. The extent to which ratios in all forms

may be profitably used in business management is fully outlined by J. H. Bliss in his "Financial and Operating Ratios in Management." As the terms "rate" and "percentage" (particularly the latter) are used extensively in the field of business, brief consideration will be given to these special types of ratios.

2. *Rate*—This term is used to represent the ratio between two numerical facts, usually over a period of time. The most familiar uses of the term in business are, perhaps, interest rate which is in the form of a percentage of the principal; and tax rates, the property tax being in the form usually of mills per dollar of assessed valuation, and the income tax being in the form of a percentage of the net income. In the field of vital statistics we have:
   (1) Death Rate—which is generally the ratio between the number of deaths and the average population, usually estimated. The *crude rate* is simply the ratio of the deaths to the population, taken customarily in thousands or hundreds of thousands; and the *refined* or *corrected rate* is the same ratio with an adjustment for the age and sex composition of the population.
   (2) Birth Rate—representing the ratio of total births to total population or to the married population.
   (3) Marriage Rate—representing generally the ratio of total marriages to total population or to those of marriageable age.
   (4) Divorce Rate—representing roughly the ratio of total divorces to total population or to the married population.
   (5) Morbidity Rates—representing roughly the ratio of sickness cases to the total population or to selected groups of the total population.
   (6) Accident Rate—representing roughly the ratio of accidents to the total population or to those exposed.
3. *Percentage*—"Percentage" is only a special type of "rate" which expresses the *number per one hundred.* When we say that 30% of the employes of a company are women we indicate that 30 per 100 are women or that out of a representative group of 100 employes, 30 are women. Or if the payroll of the company is 40% of total expenses, then $40 out of $100 expenses generally is payroll expenses. The percentage may be readily converted to the form of common or decimal frac-

tions; it is simply so many hundredths. Thus, 50% is equivalent to 50/100 or ½, or to 0.50.

"Percentage" is used much more extensively in the field of business statistics than "ratio" or "rate" and is generally more familiar. The term "rate" is used in connection with return on the investment or interest charges but in such cases it is a percentage. Any of the computations described above under "ratio" or "rate" may be readily converted to the percentage form. The fractions ½ and ¾ noted as ratios in No. 1 above may be converted to percentages simply by multiplying the respective numerators by 100 and dividing separately by the corresponding denominators. Thus, ½ is equivalent to 50% and ¾ to 75%. Likewise, a rate per 1,000 may be converted to the percentage form by making the proper computations. Thus, a death rate of 15 per 1,000 is equivalent to 1.5% and 5,000 per 100,000 is equivalent to 5%.

The treatment given here of this type of computation is necessarily brief. A more detailed discussion may be found in Chapter VI on "Ratios" in the book on "Statistics" by Bailey and Cummings.

Importance of Percentages.—The importance of ratios or percentages in statistical analysis can hardly be overemphasized. In many cases it is only by means of such computations that proper comparisons can be made. They often reveal conditions which cannot be detected by an inspection of the absolute numbers. Thus, the increase of $1.00 in the price of a pound of butter which ordinarily sells for 40 cents is a much greater increase than the increase of $5 in the price of a steel ingot which ordinarily sells for $80. In the former case there is an increase of 250% while in the latter it is only 6¼% and yet the absolute increase in the former case is only ⅕ of what it is in the latter.

This is typical of many ways in which ratios or percentages are used in practical business statistics. The net revenues of a company may be actually increasing over a period of time and yet the business may be running down—the conditions of vital importance being reflected only by percentage figures. Absences, labor turnover, tardiness and so forth are properly evaluated only

when expressed in percentage form. Constantly in the analysis of personnel, financial, and operations statistics, both internal and external, percentages are used in all sorts of ways. Quite generally the percentage or ratio form is indispensable in making important comparisons.

The Application or Use of Percentages.—Percentages are used so extensively and in such variety of cases in practical work that a complete description would be impossible even if it were advisable. Generally, percentages or ratios are computed whenever the relative value of two numbers is important. Specifically, there are certain applications which occur so frequently and are so important in statistical work as to warrant special note. In

Cotton Consumed in the United States
By Months for the Year 1923
(In thousands of running bales excluding linters)

Months	Consecutive			Cumulative		Moving Total (1)	Moving Average (1)
	Absolute Amounts	Percentages		Absolute Amounts	Percentages		
		Consecutive	Distributive				
a	b	c	d	e	f	g	h
Jan.	610	—	9.36	610	9.36	6,170	514.17
Feb.	567	93	8.70	1,177	18.07	6,265	522.08
Mar.	624	110	9.58	1,801	27.64	6,369	530.75
Apr.	577	93	8.86	2,378	36.49	6,502	541.83
May	621	108	9.52	2,999	46.03	6,628	552.33
June	542	87	8.32	3,541	54.34	6,661	555.08
July	463	85	7.11	4,004	61.45	6,666	555.50
Aug.	492	106	7.55	4,496	69.00	6,632	552.67
Sept.	484	99	7.43	4,980	76.43	6,622	551.83
Oct.	542	112	8.32	5,522	84.75	6,630	552.50
Nov.	532	98	8.16	6,054	92.91	6,583	548.58
Dec.	462	87	7.09	6,516	100.—	6,516	543.00
Total	6,516		100 —	6,516			

(Source: *Review of Economic Statistics*, Harvard Economic Service)
[1] Previous 12 months.

Table 1

this connection the classification of ratios given on pages 5-7 of the book on "Statistics" by Baily and Cummings will be of interest.

1. *Consecutive Percentages*—These are percentages of the several items in a series, each to the one immediately preceding. Column "c" in Table 1 above is illustrative.
2. *Cumulative Percentages*—This computation represents the percentages of the several cumulative items as shown in column "f" of Table 1 to the total number in the group. Table 1, page 381, will serve as an illustration of cumulative percentages in frequency series. From the latter it will be observed that the cumulation may be in either direction.
3. *Relatives*—In the development of what are called relatives, each item of a statistical series is expressed as a percentage of some single figure which is called the base. Thus in Table 1 page 452 the price for each year is expressed as a percentage of the price for 1913. More is said about relatives in the chapter on "Index Numbers," beginning on page 449.
4. *Distribution Percentages*—This computation represents the relation between a number and the several items of which it is composed. In column "c" in the frequency table referred to in No. 2 above, the frequencies in each class are expressed as percentages of the total number of frequencies in the series. Column "d" in Table 1 above is also illustrative. In similar manner the number of employes in each occupational class, each age group, or each department may be expressed as percentages of the total. Many similar distributions are also made in practical work.
5. *Miscellaneous*—There are certain other percentage computations that are used in practical business statistics which may be briefly mentioned here. Typical of these are: (1) labor turnover, which represents the ratio of force losses or replacements to the average number of employes on the payroll or average attendance; (2) mark-up percentage, which represents the percentage added by a merchant to the cost of his goods to get the selling price; (3) mark-down percentage, which represents the percentage reduction in selling price of goods; (4) stock turnover, which represents the ratio of the cost of sales for a given time period to the average inventory for the period; (5) and other computations of similar character, such as the turnover of capital and receivables.

Likewise, computations reflecting the change in certain facts over a period of time may be noted in this connection, such as the absolute and relative change (increase or decrease) over a given period of time and the absolute or relative rates of change (increase or decrease) over several time periods.[1] There is perhaps greater confusion in the use of these computations than other forms of ratios and yet the processes involved are quite simple. Assuming that the plant investment increased from $100,000 at January 1920, to $150,000 at January 1925, the computations listed are as follows:

(1) Absolute increase over the 5 years is $50,000.

(2) Relative increase on basis of January 1920, over the 5 years is $50,000 divided by $100,000, or 50%.

(3) Absolute yearly rate of increase is $50,000 divided by 5, or $10,000.

(4) Relative yearly rate of increase on a fixed base such as January 1920, is $10,000 divided by $100,000, or 10%. More precisely, the average yearly rate of increase is determined by taking the fifth root of 150% and subtracting 100 from the result, as explained in Chapter XXXIX.

It is important to note that we cannot add, subtract, average or perform other primary operations with percentages indiscriminately. Speaking generally, these operations can be properly performed only in case the factors which enter into the computations of the percentages are comparable and the bases equal. These considerations have important significance in practical work and extreme care should be exercised when such computations are undertaken. Detailed discussion of the nature of percentages with respect to these points will be found in elementary treatises on arithmetic.

[1] See "Statistical Method," by Harry Jerome, page 135.

CHAPTER XXXIV

THE METHOD OF SAMPLING

Treatment Limited in Scope.—Great economy is realized in practical statistical work through the utilization of the theory of statistical sampling. This theory enables us, through the analysis and interpretation of a relatively small number of representative data, to establish conclusions approximating those which would be indicated by the study of a very large number of the same kind of data. We are thus able conveniently to make a great many studies which otherwise would not be feasible.

There are two primary considerations in the treatment of this subject—first, the selection of the data to make up the sample, and second, the measurement of the probable error which may result from the mere fact that the data are but a sample of the entire group of data.[1] The development of the second consideration involves a refinement of mathematical processes and suggests a refinement of results which clearly do not belong in a treatise of this character. Hence the present discussion will be confined to a consideration of the method of selecting the sample data and of rough practical tests of their representativeness.

The Meaning of Statistical Sampling.—A statistical sample is a miniature picture of the entire group of the particular type or kind of data of which the sample is a part. Thus we may have a sample group of such facts as wage data, absence data, sales data, purchasing data, or production data.

Studying a relatively small number of items does not necessarily satisfy the conditions of sampling; the data chosen must be *representative* of the larger group of which they are a part. To be representative the sample must contain the several varieties of data found in the larger group and in approximately the same propor-

[1] The group of data of which the sample is a representative portion is designated conventionally by the terms "universe" or "population." In a study of rates of wages of unskilled workers in the United States, for example, during the month of January 1925, the "universe" or "population" would be the entire group of wage payments made to this class of workers.

tions. If, for example, 100,000 price data are distributed as follows:

20	cents	– 10,000
25	"	– 20,000
30	"	– 30,000
35	"	– 25,000
40	"	– 15,000

A sample of this group to be perfectly representative must contain a proportional number of data in each price group. Of the following samples, Group A is representative and Group B is not:

Group A			*Group B*		
20	cents	– 10	20	cents	– 50
25	"	– 20	25	"	– 20
30	"	– 30	30	"	– 10
35	"	– 25	35	"	– 5
40	"	– 15	40	"	– 30

The point is that the study of a relatively small group of data will reveal the same conclusion that the study of a very large group of the same kind of data will reveal, providing the sample is truly representative.

In the more technical sense, however, statistical sampling has to do with the method of determining a representative sample in cases where the relative proportion of different kinds or sizes of data is not known and of evaluating the representativeness of the sample. The accuracy of the results obtained from the analysis of the sample increases with the representativeness of the sample and, as indicated above, if the sample is to give us results identical with those obtained from all the cases, it must be perfectly representative. If, however, the sample is only approximately representative the results will generally be sufficiently accurate for practical purposes.

It should be noted also that the selection of the sample data is not necessarily tied up with the collection process: the collection process is not materially affected by it.

In any case, we have to determine, before the collection is started, just what data are to be collected. This is true for complete surveys as well as sample studies. On the other hand, even if all the data are on the systematic records of the statistical department we still have to decide the number and character that will be required in the particular case. In other words, the fact that

we are not obliged to *collect* the data does not eliminate nor seriously alter the problem of sampling.

The Ideal Sample.—The ideal sample is one that is perfectly representative. That is, it is one in which the several kinds of data are represented identically in the same proportion as in the larger group. Suppose, for example, that we wish to determine the average rate of pay of 200,000 employes distributed as follows:

80,000	receiving	$15	per	week
20,000	"	18	"	"
40,000	"	20	"	"
20,000	"	23	"	"
30,000	"	25	"	"
10,000	"	30	"	"

It is obvious that we obtain the same results from the following in which only 20 employes are involved that would result from a computation based on the entire 200,000:

$$
\begin{array}{rl}
8 \times 15 = & 120 \\
2 \times 18 = & 36 \\
4 \times 20 = & 80 \\
2 \times 23 = & 46 \\
3 \times 25 = & 75 \\
1 \times 30 = & 30 \\
\hline
20 & 387 \\
 & \$19.35
\end{array}
$$

The different wage groups are represented in the sample in exactly the same proportion in which they appear in the group of 200,000. In general we can obtain the same results from the ideal sample, however small, that can be obtained from a very large group of facts.

The Sample in Practical Work.—The ideal sample is rare in practical work. Generally, we do not know the relative proportions of the different kinds of data and have no means of finding them out. Consequently, samples in practical work only approach the character of the ideal.

Furthermore, in most cases in practical work we have no means of determining exactly how nearly the ideal the sample is. For example, we may wish to study the time required to produce a unit of a particular class of factory product in the United States. There may be millions of units of this particular article produced

and no one has ever attempted to measure the time required to produce each separate unit. That is, we have no means of knowing precisely what the proportional distribution of the different periods is and, hence, we cannot measure the distribution of any sample against the distribution of the entire group. This is generally true for all kinds of data in practical work.

Fortunately, however, such a measurement is not necessary. We can accomplish the same results sufficiently accurately for practical purposes in another way: namely, through the use of the law of probability.

The Theory of Probability.—The theory of probability as applied to statistics means that a reasonably large number of data chosen *at random* from among a very large group are almost certain on *the average* to have the characteristics of the group,[2] *i.e.*, that the sample group contains approximately the proper proportion of each type of data. In other words, our aim is to get a sample in which the distribution approaches that of the ideal sample and through lack of ability exactly to select the relative number of each type of case, we employ the law of probability.

The theory of probability in its simpler aspects is illustrated in many ways. For example, if two balls, one black and one white, are placed in a box and a boy who is blindfolded draws one ball, the chance that he will draw a black ball is one out of two or "50-50" or more technically ½. Likewise, if twenty black balls and twenty white balls are placed in the box and the boy draws four balls under similar circumstances, it is more likely that he will draw two of each color than any other combination. The mathematical probability that this or any other possible combination will result from the draw, can be computed with preciseness. But further development of mathematical probability would go beyond the scope of the present work. It is sufficient to note the general principal that in this illustration if each ball is given an equal chance to be drawn, the number of any one kind actually drawn will tend to be proportional to the number of this kind in the box.

Applying this fundamental principle to statistics, if a group of

[2] W. I. King, "The Elements of Statistical Method," page 28.

8,000 sales data is made up of 1,000 of one grade of goods, 2,000 of another and 5,000 of another, and we select 800 from this group in such a way as to give each one an equal chance to be selected, the 800 selected would likely contain a substantially proportional number of each group, distributed in the same general proportion as those in the main group. Likewise, if we study a reasonably large number of representative family budgets in a certain economic class, or wages of factory workers in a reasonably large number of representative establishments, or a reasonably large number of labor turnover facts, dividend rates, and so forth, the results for practical purposes will be the same as if we should increase the number in each case to an indefinitely large group.

This is what is technically meant by statistical sampling. Conscious selection plays little or no part in this process. However, in most cases in practical work conscious selection does enter in to a certain extent and we shall, therefore, broaden the conception of sampling to include samples chosen through *conscious selection* as well as those chosen through *random selection.*

The Practical Procedure.—It is clear from what has been said that there are two principal methods of selecting the sample in practical work: conscious selection, in which we make little or no use of probability, and random selection, in which we depend on the law of probability to give us the proper distribution of the data.

Conscious Selection.—In conscious selection we survey the different classes or types of items and consciously select a group which is satisfactorily typical of all the facts. This method may be illustrated by the procedure followed in developing wage rates for the community. The particular companies in the community to be included in the study are determined entirely by conscious selection. We evaluate the several companies for the purpose at hand and choose those which in our judgment will be representative of the entire group. Likewise, in developing index numbers of wholesale prices or of the cost of living, the commodities to be included are chosen through conscious selection. This method is used extensively in practical work in all three groups of facts—personnel, financial and operations.

Random Selection.—In many cases, however, we are unable to make conscious selection of the items that are to be included in the sample. In such cases we resort to random selection. That is, we select the required number of cases at random from the entire group of data (the "universe") and depend on the law of probability to make them representative. By at random we mean in general that each item in the entire group is given an equal chance to be included in the sample. It is essential that this element of the process be observed, since the law of probability will operate to produce proper results only when the cases are selected at random. It is important to note as a point that is frequently overlooked that this requirement implies that the selection of one item will not alter the chances of another's being selected. Typical of the application of this method in practical work is the selection of price data in developing index numbers of wholesale prices, retail prices, and cost of living.

Many practical suggestions may be offered for securing random selection in certain kinds of problems. Numbers may be assigned to the separate items in the group from which the sample is to be selected. These numbers may then be placed in a box, well mixed up together, and drawn out after the fashion described in the illustration of the black and white balls above. Or, they may be arranged in order and selected in some regular manner, say, every 5th, 10th or 15th. Thus in a sample study of contracts, say, in a large public utility where hundreds of thousands of contracts are filed alphabetically, the sample may be selected by drawing one in about every 1,000, 10,000, or 100,000 or even one from each file case. Again if these contracts are numbered the randomness of the sample may be secured by drawing numbers as suggested above.

The Size of the Sample.—In practical work there is always danger of incorrect sampling. The selection of the sample data may be consciously biased or, in random selection, misrepresentation may result from some factor unknown to us. Hence, in developing sample studies ourselves or in examining those made by others, we should make sure that the sample is "good," that is, one which contains at least approximately a proportional number

of each kind of data found in the group of which it is a part. This proper distribution of items in the sample is determined in conscious selection through the judgment of the statistician; in random selection we accomplish a proper proportion of items through the law of probability.

In the latter case, as we do not know the distribution of the facts in the group from which we are taking a sample, we must devise some means of determining whether the sample is "good." In this connection, the number of data in the sample becomes important. As the representativeness of the sample increases with the number of items it contains, we get a better sample by increasing the number of data. There is a point, however, in each problem at which the representativeness of the sample approaches the ideal sufficiently for practical purposes and which makes additional data unnecessary. This point is determined practically through the trial and error process. We select several samples each containing what seems to be a proper number of data and perform some simple computation, say, the average. If the variations in the results are substantial we enlarge the number of data in a second series of samples and perform the same computation. When the result of this computation remains approximately constant among the several samples, *i. e.,* when we have reached what is sometimes called the *point of stability*—we assume the sample is "good." As it is more economical in time and expense to handle a fewer number of data, it is advisable to select the smallest number which will be properly representative.

This test can be applied, of course, only when the data are on hand. In determining in advance of collection the amount of data that will be required, we have to depend on our general knowledge, a survey of the known facts, an examination of similar studies, and other practical means. Even where the test can be applied we should check the sample in these commonsense ways.

Mathematical Tests.—As stated at the beginning of this chapter more refined measurements of the representativeness of the sample can be accomplished through certain mathematical computations, chiefly the computation of the *probable error.* Thus,

we may compute the *probable error* of the representativeness of the sample itself or of certain computations derived from it such as the probable error of the means, the standard deviation, the coefficient of correlation, and other like computations. These processes, however, are not pertinent to a discussion of this kind. They are discussed at length in the references given below.

Conclusion.—While sampling is very serviceable in the selection of data in practical work and should be used whenever possible, there are problems to which it does not generally apply. Typical of these are internal studies relative to the amount of revenues and expenses, sales, output, advertising, absences, and labor turnover. Likewise, problems involving the number of gainfully employed in the community, the number of potential employes released by the schools and colleges, and the number of immigrant workers coming into the country can be accurately determined only by making an actual count. Such problems, sometimes referred to as complete *inductions,* do not lend themselves to the method of sampling.

In other problems such as the analysis of retail prices, wages, or sales in the community where the sampling process is either necessary or highly advisable, we go as far as we can through conscious selection and then resort to random selection and the law of probability. In such cases, the sample will not yield identically the same results as all the data except when it is ideal and rarely do we have the ideal sample in practical work. However, the methods outlined in this chapter will generally produce samples sufficiently "good" for practical purposes.

References

Bowley, Arthur L. Elements of Statistics, 2nd Ed., 1902, pp. 308-316.
—— An Elementary Manual of Statistics, 1915, Ch. VII.
Jerome, Harry. Statistical Method, 1924, Chs. II, X.
Mills, Frederick C. Statistical Methods, 1924, Ch. XVI.
Rietz, H. L. Handbook of Mathematical Statistics, 1924, Ch. V.
Secrist, Horace. An Introduction to Statistical Methods, 1917, pp. 20-22.
—— Readings and Problems in Statistical Methods, 1920, pp. 62-64 (Sampling of Coal), pp. 91-110 (Alternative to a Count), pp. 111-124 (Development of Markets).
Yule, G. U. An Introduction to the Theory of Statistics, 1919, Part III.

CHAPTER XXXV

FREQUENCY CLASSIFICATION

Meaning of Frequency Classification.—Further treatment of quantitative classification has to do primarily with the formation of magnitude classes. Broadly considered, quantitative data are the result of individual quantitative observations expressed in units of *measurement,* distinction being made between this and the quantitative aspect of all statistical data which results from the process of counting. The units of measurement will vary greatly in practical work. Thus we may measure the ages of employes, their periods of service, and the lengths of absences and tardinesses, in units of time; the size of farms, the mileage of tires, and freight haulage, in units of space; interest rates, individual incomes, security yields, mortgage loans, prices, and wages, in units of value; daily sales, daily deliveries, and daily completions, in units of product; and so forth. Now, quantitative classification as here understood consists merely in grouping these individual observations into magnitude classes in order better to reveal their essential characteristics. In accomplishing this grouping we may form the classes before the data are grouped and then place the items in the classes, or we may take the group of data and divide it into equal classes by means of quartiles (values which divide the series into four equal parts), deciles (values which divide the series into ten equal parts), and percentiles (values which divide the series into one hundred equal parts), and so forth. It is the former that we have in mind here. It will be noted that the data may be the result of taking single measurements of many different units, or many measurements of a single unit, although in elementary business statistics we have the former type almost exclusively.

Because this type of classification really reflects the number of times individual measurements or groups of measurements occur—*i. e.,* their frequency—it is known as a frequency distri-

bution.[1] In tabular form the frequency distribution becomes the *frequency table* which may be defined as an orderly association of a series of measurements and the frequencies with which they occur. Such a table is constructed by listing singly or in groups and in ascending or descending order the units of measurement, and by arranging opposite them their corresponding frequencies.

Since we generally group data at different time or space periods historically or spatially, there is danger of the beginner's getting the impression that these are the only methods of classification in such cases. It should be noted, however, that such data may also be grouped into frequency classes. It is simply a question of making the quantitative value of the data the basis of classification instead of the other factors. Frequency analysis of the data for different time or space periods will reveal the number of times the several items of the series occurred during the entire time period or within the entire territory.

The Frequency Curve.—The graphic representation of the frequency series is usually effected by means of the curve, although a series of vertical bars may be used. The principles and practices outlined in Chapter XXIX, beginning on page 280, for the construction of the historical curve will apply generally to the construction of the frequency curve; the relative values of the two scales, the location of the points, the connection of the points, the methods of smoothing, and so forth, will be determined generally in the same way.

Figure 101, below, will clarify the procedure to be followed in constructing the frequency curve. It will be noted that the measurements are scaled horizontally and the frequencies vertically. In plotting the grouped series, the general assumption is followed of taking the mid-point of the group as representative of the group. It is important to note also, as stated on page 388, that any variation in the value of the class-interval must be taken into account. Equal distances on the horizontal scale must be given to equal units of measurement and if the class interval is increased the distance on the scale must be increased accordingly. Likewise, the distance in the vertical scale must be adjusted. If,

[1] The term "frequency distribution" in the present discussion will be confined to the narrower conception of the classification of magnitude classes.

for example, the class interval and, consequently, the distance allowed to the class on the horizontal scale, be doubled, the distance on the vertical scale must be halved. Or, to state the same thing in another way, we should represent the larger class group by reducing the data contained in it to an average for the regular sized group.

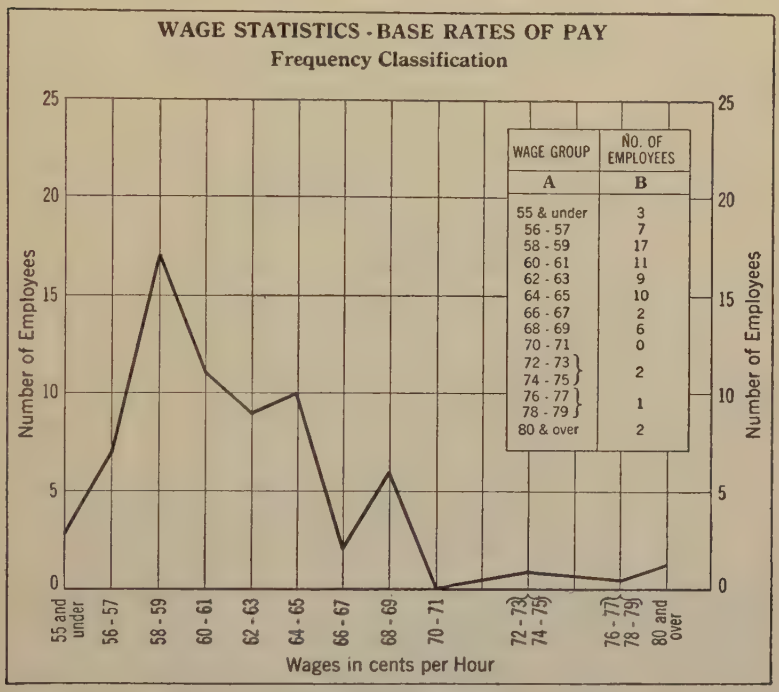


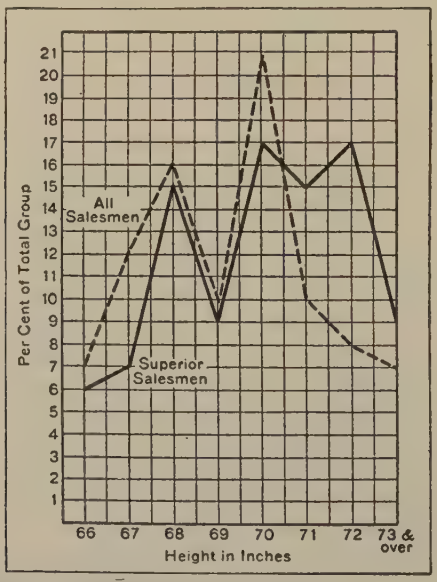
FIGURE 101.—Illustrating the Method of Plotting Frequency Data Where the Groups are Unequal

The two groups next to the last group are double in width and consequently the value of frequencies is halved to locate the point.

We may also have the frequency band curve in which component parts of a frequency series are plotted on the same grid as the total such as total labor losses by wage groups and losses for separate period of service groups. The construction is similar to that of the historical band curve shown in Figure 68, page 320.

Continuous and Discontinuous Series.—From the standpoint of classification two broad types of series are distinguished as follows:

CONTINUOUS SERIES.—A continuous series is one in which the measurements cannot be accurately determined and may fall at any point within the range of the series, according to the degree



(From O. R. Johnson in *Administration*, June 1923)

FIGURE 102.—Comparison According to Height of 506 Salesmen with 54 Superior Salesmen, Showing Tendency of the Latter to be of Unusual Height

A continuous frequency series

of refinement in the unit of measurement. This is illustrated by the data showing the height of salesmen given in Table 1 below and Figure 102 above. The measurements recorded in this table were arbitrarily taken to the nearest inch and the classes based on them are likewise arbitrary. To say that there were no men in the group who measured between 68 and 69 inches, for example, is incorrect. If we should make the measurement, say, to the nearest 10th of an inch, we should probably find several measurements between 68 and 69 inches. The point is that a

measurement may fall at any point within the range depending on the minuteness of the measurement—*i. e.*, the measurements may run continuously from one extreme to the other if an infinite number of items be measured and the unit of measurement be made infinitely small.

HEIGHT OF A GROUP OF SALESMEN

To the Nearest Inch

Height in Inches	No. of Salesmen		Cumulative—Less Than		Cumulative—More Than	
	Absolute	Per Cent of Total	Absolute	Per Cent of Total	Absolute	Per Cent of Total
a	b	c	d	e	f	g
65 and under	40	8	40	8	505	100
66	35	7	75	15	465	92
67	61	12	136	27	430	85
68	81	16	217	43	369	73
69	51	10	268	53	288	57
70	106	21	374	74	237	47
71	51	10	425	84	131	26
72	40	8	465	92	80	16
73 and over	40	8	505	100	40	8
Total	505	100				

TABLE I.—Heights to the Nearest Inch of a Group of 505 Salesmen Hired by A. M. Collins Manufacturing Company, as reported by O. R. Johnson in *Administration*

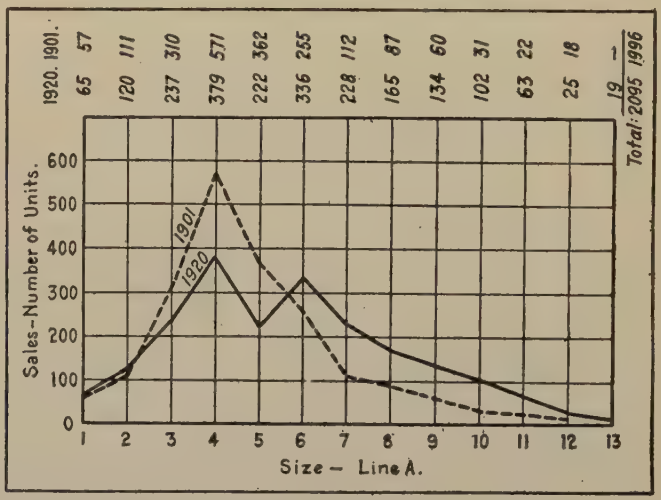
The figures have been changed slightly to serve the particular purpose of this illustration. It is assumed that each entry contains all items from ½ inch smaller to ½ inch larger—e. g., 68 includes all from 67.5 to 68.5.

DISCONTINUOUS SERIES.—A discontinuous series is one in which the minuteness of the measurements is limited by the character of the statistical unit. This type is illustrated by the number of sales for each size of article as shown in Figure 103 below and the accompanying table: we can never obtain measurement between the sizes of the articles.

Similar limitations obtain in a count of the number of employes, production units, deliveries, absences, and so forth. In

certain other types of data the discontinuous character is arbitrarily imposed. If wage rates, for example, are on a weekly basis and are held only to the nearest dollar, then we should not have a measurement smaller than a dollar, which becomes a definite limitation on the minuteness of the measurement. Many series in practical work are of the discontinuous type.

The primary significance of the distinction between continuous and discontinuous series lies in the fact that in the former, if the series is representative, we can by the application of more refined



(From article by David B. Porter in *Management and Administration*, Mar. 1924)

FIGURE 103.—Number of Sales by Sizes of Articles

A discontinuous frequency series.

statistical methods obtain a general picture of the series which would result from an infinite number of items and from this derived distribution we can interpolate for refinements between recorded measurements.

Such interpolation cannot be made with the discontinuous series. The measurements cannot be taken to a finer degree than the statistical unit and no number of measurements however great would change this fact. The spaces between the measurements in Figure 103 above, for example, has no significance other than to aid the eye in following the outline made by the recorded data: it is impossible for measurements to fall in these spaces.

HOURLY RATES OF PAY—MEN FACTORY WORKERS

Period of Service, 3 to 7 Years
(In cents per hour)

64	59	57	67	61	72	78	69	62
58	62	62	59	59	57	54	61	59
56	56	80	65	57	64	59	56	64
55	60	59	57	63	59	66	62	61
60	61	61	62	79	57	59	69	59
51	56	79	58	59	72	56	67	52
48	59	59	54	57	64	69	57	68
60	62	57	62	63	59	78	62	59
58	60	62	59	56	57	79	56	62
48	59	61	61	64	56	62	68	55
63	62	74	66	66	64	59	59	68
60	83	59	57	59	59	69	55	54
54	61	72	62	63	56	72	57	67
58	70	57	54	74	61	63	67	61
58	60	61	61	64	57	74	62	59
53	61	65	59	67	56	59	52	55
50	80	59	63	59	54	62	68	62
58	60	62	61	64	64	69	56	64
60	65	57	74	64	59	54	59	57
54	59	88	64	59	64	56	64	74
63	61	61	54	72	61	69	59	62
50	66	57	56	54	72	57	61	56
53	59	84	61	64	64	64	67	55
56	59	59	78	75	78	74	52	59
59	81	77	63	57	61	63	57	57

TABLE 2.—This table illustrates the unsystematic arrangement of statistical data in their original form

Steps in Analyzing Frequency Series.—It will be observed from the table accompanying Figure 101 above, that the frequency distribution consists primarily of, first, a series of measurements, and, second, a series of corresponding frequencies. In its development the following main steps are observed:

1. *Array the Data*—In which the original observations are arranged in order of size, ascending or descending.
2. *Select the Classes*—In which the number of classes and their widths are determined. The groups into which the items are arranged are called class-groups or classes.
3. *Determine the Class Limits*—The class limits are the bound-

Hourly Rates of Pay (Cents per hour)	Number of Employes
a	b
48	2
49	0
50	2
51	1
52	3
53	2
54	10
55	5
56	15
57	20
58	6
59	36
60	8
61	20
62	18
63	9
64	17
65	3
66	4
67	6
68	4
69	6
70	1
71	0
72	6
73	0
74	6
75	1
76	0
77	1
78	4
79	3
80	2
81	1
82	0
83	1
84	1
85	0
86	0
87	0
88	1
	225

TABLE 3.—A simple array, showing the data of Table 2 arranged in groups with one cent as class interval.

ary lines of the class groups. The size of the class or the distance between the class limits is called the class interval. This step involves the specific manner of expressing the class limits.

4. *Enter the Items*—This step consists merely in placing the individual items in their respective classes.

These steps will be discussed in order.

The Statistical Array.—The first step in frequency classification is the formation of the statistical array. This process consists of arranging the measurements in order of size, ascending or descending, and associating with them the corresponding frequencies. See Tables 2, 3, and 4.

Table 2 shows the measurements in the order in which they were recorded at the time of collection. In arraying raw data we first develop the primary array in which the separate individual measurements are arranged in order of size and from this primary array we proceed to further classification by grouping the items. Tables 4 and 5 illustrate grouped arrays.

The detailed method of procedure in forming the array will vary according to

M	Tally	Count
14	////	4
15	𝍸 //	7
16	𝍸𝍸///	13
17	𝍸/	6
18	///	3
19	//	2

TABLE 3a

the character of the data and the ingenuity of the statistician. The data may be organized by hand in some such way as that shown in Table 3a.

This method consists simply of tallying the frequencies against the measurements and counting the tallies. It is quite serviceable where the number of data are few but where the data run into large numbers it is both cumbersome and inaccurate. The work cannot be checked except by repetition and this is expensive in both time and money.

Hourly Rates of Pay (In two-cent groups)	Number of Employes
a	b
48 – 49	2
50 – 51	3
52 – 53	5
54 – 55	15
56 – 57	35
58 – 59	42
60 – 61	28
62 – 63	27
64 – 65	20
66 – 67	10
68 – 69	10
70 – 71	1
72 – 73	6
74 – 75	7
76 – 77	1
78 – 79	7
80 – 81	3
82 – 83	1
84 – 85	1
86 – 87	0
88 – 89	1
	225

TABLE 4.—The grouped array showing data of Table 3, arranged in two cent groups.

Hourly Rates of Pay (In five-cent groups)	Number of Employes
a	b
46 – 50	4
51 – 55	21
56 – 60	85
61 – 65	67
66 – 70	21
71 – 75	13
76 – 80	10
81 – 85	3
86 – 89	1
	225

TABLE 5.—Grouped array, showing data of Table 3 arranged in five-cent groups.

When the number of items is large it is preferable to transfer them to the Hollerith cards. In this form they may be readily classified in all possible ways. The method of using the Hollerith system and the service that may be derived from it are fully developed in Chapter XLV, beginning on page 493.

The use of a separate card for each individual item is sometimes feasible. In this case the cards are grouped and counted. The use of this device likewise will be confined for the most part to situations in which the number of data are comparatively few.

Other methods will be devised to meet the peculiar requirements of the particular situation. While the data may be arrayed in either ascending or descending order, the former is generally preferable in the fields of business statistics. The method of procedure as set forth above applies without distinction to continuous and discontinuous series.

Selecting the Classes.—From the material as arranged in the primary statistical array, we proceed to the second step which consists of selecting the classes or groups in which we wish the data finally to be organized. These classes may be determined by certain characteristics pointed out below but usually they must be established arbitrarily. The size of the class will vary from very small in which the class interval is unity or less to the larger classes in which the class interval may be several times unity. In this connection, it must be kept in mind that the position of the classes is quite as important as their size.

In further analysis of the data or in otherwise making use of the series, we assume that the items of each class are equally distributed over the class and are represented by the mid-value of the class. This assumption is made in computing averages as noted in Chapters XXXVI to XXXVIII, in plotting the frequency graph as shown in Figure 101 on page 379, and in other uses generally to which the series may be put.

To aid the judgment in selecting the classes there are certain general principles that apply in the absence of any specific facts which determine the classes to use:

Minimum Number of Classes—From the standpoint of economy in the further analysis of the data the number of classes should be reduced to a minimum. This means that the classes should be as large as is consistent with the requirements of the situation. Some have suggested that the number of classes should range somewhere within the limits of 15 to 25 and that we could simply take the difference between the lowest and highest items in the group and divide by the number of classes to get the class interval. Such a procedure, however, is wholly unscientific: it fits the facts of the series to a purely arbitrary classification rather than fitting the classification to the facts. However, experience in handling certain kinds of data may enable us to figure on a limited number of classes.

2. *Required Degree of Accuracy*—On the other hand, as the classes are enlarged, detailed accuracy is sacrificed. In the grouped series we treat all the items in a single class-group as equal to the mid-value of the class-group and as the classes are enlarged the error from this assumption increases. In any given case, a sufficient degree of accuracy must be maintained to serve the specific purpose of the analysis else the entire development loses its value. Enlarging the class interval smooths the series and makes for economy but we must not unduly sacrifice accuracy to these ends.

Supplementing these two fundamental principles is a third which prescribes the necessity of having the distribution reasonably smooth. This is necessary in order that the general character of the series may be revealed.

The operation of these principles necessitates at once, (1) the determination of the allowable difference between the degree of accuracy of the condensed series and that of the original detail, and (2) a means of measuring this difference.

The first of these will vary, the error that may be allowed being determined by the purpose for which the results are to be used.

The second requirement may be satisfied in a number of ways. The weighted average is a simple test. If the weighted average computed from the condensed series is far removed from that taken from the primary array, the classes should be reduced in

size or the condensed series used with the greatest caution. The method then really becomes a trial and error process. We select one grouping and test for accuracy; and then another, and another. When we have reached the largest class interval that will still give satisfactory accuracy, we have the proper grouping.

A second test may take the form of definitely checking up how near the items in the group actually coincide with the mid-value. This can be done directly from the detail if we are arranging a series of data into larger classes: if we are determining a classification preparatory to collection, the test will have to be made on the basis of a sample or by inspection of similar studies. A comparison of Tables 4 and 5 above with Table 3 above will illustrate this procedure. As a matter of fact we soon establish the proper classes for different kinds of data in practical work.

Generally, the class intervals of a series should be equal. Violation of this practice is apt to lead to error and may impose obvious limitations on the use of the data. However, it is proper to vary the class interval if the particular data call for it but in such case the change in the value of the class interval should be clearly indicated by a bracket or by some other device.

In applying the above principles to discontinuous series, we must bear in mind that the character of the data generally imposes a limitation on the refinement of the classification, to disregard which leads to serious error.

Ease and accuracy of classification and subsequent computation should also be taken into account in selecting the classes. Since each class-group is represented by its mid-point, it is advisable to select the classes so that the mid-point will be an integer free from fractions, if such a classification is consistent with other factors in the situation.

Classes Imposed by Nature of Data and Purpose.—There are conditions, however, where the nature of the data and the purpose of the classification will determine the exact delineation of the classes. When the items tend to cluster about certain points in the series, for example, as in interest, dividend, and mortgage rates, and age statistics, these points should be made the mid-points of the classes. Inspection should be made, of course, to ascertain

whether the items are fairly symmetrically distributed from these concentration points and the practical tests should be applied. If, in the primary array most of the items are actually concentrated at the mid-points chosen, the assumption that the items in a group are equal to the mid-point is well established. There is little difference between the average obtained from such a grouping and that obtained from the primary array.

Furthermore, special uses of the data frequently require that the classification provide certain specific groupings. Suppose, for example, that the vacation allowance of the company is graduated on the basis of period of service as follows:

Less than one year—No vacation.
One year but less than two—One week.
Two years but less than five—Two weeks.
Five years but less than fifteen—Three weeks.
Fifteen years and over—Four weeks.

In classifying employes by periods of service, the classes should be selected to conform to these provisions—*i.e.*, we should be able to get from the classification the number of employes who have a period of service of "Less than one year," "One year but less than two," and so forth. Similar problems will arise in classifying absence or labor turnover data to permit comparison with outside concerns or other standards. In fact, wherever comparisons are involved the class grouping must be consistent in the different series being compared.

The Class Limits.—There are two broad methods of describing the classes into which the data are to be grouped:

1. *By Indicating the Class Limits*—There are four different ways in which the class limits may be expressed. These are illustrated in Tables 6 and 7 below. The classes in Table 6 are generally intended to express identical class limits but different from those in Table 7 which are also intended to be identical. The methods of expression, however, in corresponding columns of the two tables are the same.

   The efficacy of these methods depends primarily on the definiteness with which they indicate the class limits. From this point of view the method of column "a" in Table 6 and

column "e" in Table 7, is unquestionably preferable. The others are ambiguous. In column "b," for example, the last class is generally intended to include only those items which measure exactly *20 or above* but it may quite as properly be interpreted to include all items above 19.5. Similar interpretation might be placed on other classes of this column and of those in columns, "c," "d," "f," "g" and "h." It is

Column a	Column b	Column c	Column d
0– 4.99	0– 5	0 under 5	0– 4
5– 9.99	5–10	5 " 10	5– 9
10–14.99	10–15	10 " 15	10–14
15–19.99	15–20	15 " 20	15–19
20 and over	20 and over	20 and over	20 and over

TABLE 6

obvious that the type of classification indicated in columns "a" and "e" would be possible only in case special provision were made for it in the collection of the data.

In the discontinuous series, the class limits are obviously more readily expressed since there are no measurements smaller than unity. In column "f" of Table 7 below, for example, the class limits in a discontinuous series could

Column e	Column f	Column g	Column h
0–5	0– 5	5 and under	0– 5
5.01–10	5–10	Over 5 to 10	6–10
10.01–15	10–15	" 10 " 15	11–15
15.01–20	15–20	" 15 " 20	16–20
Over 20	Over 20	Over 20	Over 20

TABLE 7

mean only one thing—*i. e.*, that the upper limit of the class is included. A measurement of 20 is included in the next to the last class and 21 in the last one, and as there are no measurements between them, the grouping is clear.

2. *By Indicating the Mid-values*—This method consists merely of stating the mid-points of the classes. It is illustrated in Table 8 below, columns "i" and "j" representing classifica-

tions in which the class intervals are *unity* and *greater than unity respectively.*

Generally, the classes so indicated are intended to include an equal number of items on either side. In column "i," for example, the first class, unless otherwise specified, is meant to include all items measuring from 9.5 to 10.5 and the first class in column "j" everything from 0 to 5. They may, however, be interpreted in any one of the different ways noted in No. 1 above. For example, the first class in column "j" may include (1) all the items from 0 to 4.99

(Class Interval Unity)	Classes Indicated by Mid-Points
Column i	Column j
10	2.5
11	7.5
12	12.5
13	17.5
14	22.5
15	27.5
16	32.5
17	37.5

TABLE 8

as indicated in column "a" above, (2) all items from 0 up to and including 5 as indicated in column "e" above, or (3) all items from 0 to 4.5 a possible interpretation of columns "b," "c," and "d" or from 0 to 5.5 a possible interpretation of columns "f," "g," and "h."

Here, again the discontinuous series requires special consideration and the nature of the mid-value must be determined in the light of the character of the data themselves. The mid-value must always be an actual case.

The foregoing has been presented primarily for the purpose of indicating the possibility of confusion in the interpretation of the class limits and the necessity of making them absolutely clear. It is evident from what has been said that the statement of the class limits should be *fool-proof* both for use in the collection process

and in further classification after the data have been collected. Failure to do this may lead to classification which is inconsistent and otherwise improper and also to incorrect interpretations and hence false conclusions from classifications that have been made.

Satisfactory statement of the class limits may be made by the methods of columns "a" and "e" on page 390, or by using any one of the other methods and supplementing it with an explanatory note definitely indicating the specific meaning of the expression. This is the conservative procedure. We must first determine the general type of grouping—that illustrated by Table 6 above, or by Table 7—and then choose the preferable method of expressing the class limits to indicate the grouping chosen. The type of grouping may have to be decided before the collection of the data. The type shown in Table 8 above, column "j," cannot be made from data which are collected, for example, so that all measurements from 4.5 to 5.5 are recorded as 5.

Entering the Data.—In entering the data, the primary consideration is that the individual items be included in the proper classes. For the most part, this requires only the usual caution necessary in handling statistical data.

Where an individual measurement is the same size as the class limits, the measurement may be figured to a smaller fraction if possible, when its proper classification will be apparent. If this cannot be done, the items may be distributed, half to one class and half to the other, or it may be divided on the basis of the number of items in the two adjacent classes. As a matter of practice, few cases of this kind arise and they may usually be disposed of in the most convenient way without serious error. It should be noted incidentally that a zero should be entered opposite classes which have no frequencies; in no case, should the classes be omitted because they have no frequencies.

Form of Frequency Series.—The form of the frequency series will be considered from two points of view:

1. The general character of the distribution on which basis we have a variety of forms the most frequent of which in practical business statistics is the bell-shaped distribution.

2. The organization of the data contained in the series on which basis we have principally simple and cumulative.

These different forms will be given only brief consideration: more extensive treatment will be found in the references listed on page 401.

METAL TRADES EMPLOYES

Hourly Rates of Pay

Earnings per hour (cents)	Simple Frequency	Cumulative	
		Less Than	More Than
a	b	c	d
21 – 24	2	2	2,842
25 – 28	5	7	2,840
29 – 32	29	36	2,835
33 – 36	44	80	2,806
37 – 40	140	220	2,762
41 – 44	107	327	2,622
45 – 48	165	492	2,515
49 – 52	210	702	2,350
53 – 56	262	964	2,140
57 – 60	307	1,271	1,878
61 – 64	177	1,448	1,571
65 – 68	303	1,751	1,394
69 – 72	269	2,020	1,091
73 – 76	292	2,312	822
77 – 80	271	2,583	530
81 – 84	52	2,635	259
85 – 88	103	2,738	207
89 – 92	59	2,797	104
93 – 96	15	2,812	45
97 – 100	16	2,828	30
101 – 104	10	2,838	14
105 – 108	1	2,839	4
109 – 112	2	2,841	3
113 – 116	0	2,841	1
117 – 120	1	2,842	1

(Source: National Metal Trades Association—N. Y. & N. J. Branch)

TABLE 9.—Simple and Cumulative Frequency Series (Data for Figure 108)

The Bell-Shaped Distribution.—Most of the series in practical business statistics take, roughly, a bell-like shape when ex-

pressed in graphic form. They will vary from the moderately asymmetrical, as shown in Figure 102 above, to the widely skewed series illustrated by Figure 105 given below. Some distributions, however, do not show concentration points of any particular

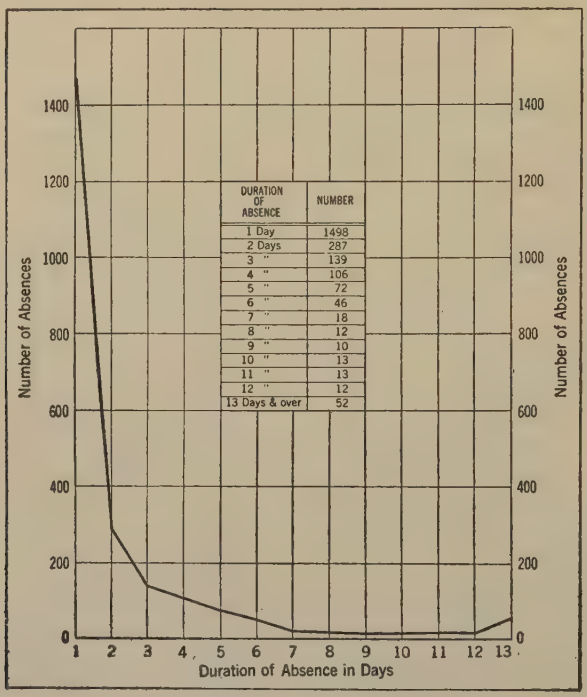


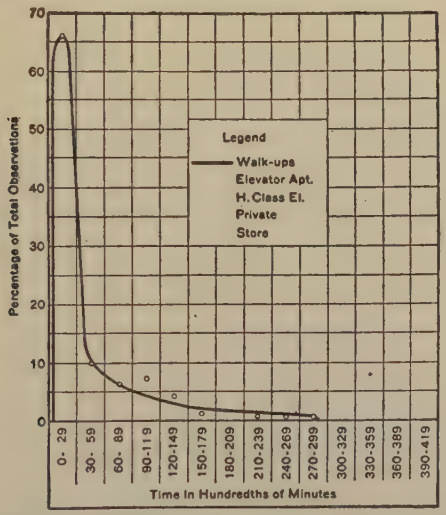
FIGURE 104.—The J-Shaped Distribution

significance, and others show more than one concentration point (Table 3, page 426), and Figure 110, page 425. The latter is generally known as bi-modal, tri-modal, and so forth according to the number of concentration points.

Miscellaneous Forms.—Certain groups of data with which we deal frequently in practical work take a form different from the bell-shape. In the analysis of absences by duration of absence, for example, we get a distribution shown in Figure 104 above in which the data are concentrated toward the lower

limit of the scale, resulting in a distribution the curve-graph of which is sometimes referred to as a J-shaped curve. Figure 105 representing data in the field of operations is of the same general form. Distribution of incomes as shown by income tax statistics, accidents by periods of service, labor losses by period of service, and other groups will reveal the same general form of distribution.

On the other hand, the analysis of mortality or accident statistics in the general population by age groups will show a relative



(From H. B. Wess in *Administration*, Mar. 1923)

FIGURE 105.—Distribution of Totals of all Observations on "C" Operation The J-shaped distribution.

concentration at both the lower and upper limits of the scale, the curve-graph of which will take the form sometimes referred to as the U-shaped curve. This form is not met with so frequently in elementary business statistics as the distribution mentioned above.

The important point is that all distributions are not of the bell-shaped form. We must be prepared to analyze whatever facts present themselves and in no case should we attempt to *force* a series into the bell-shaped form.

The Simple Series.—The simple association of the items with

their corresponding class groups constitutes the simple frequency series. The items may be expressed in absolute form or as percentages of the total number of items. Table 1, page 381, is illustrative. It is frequently difficult to compare different series in absolute numbers because of the difference in the number of items and consequently we usually convert the data to percentage form

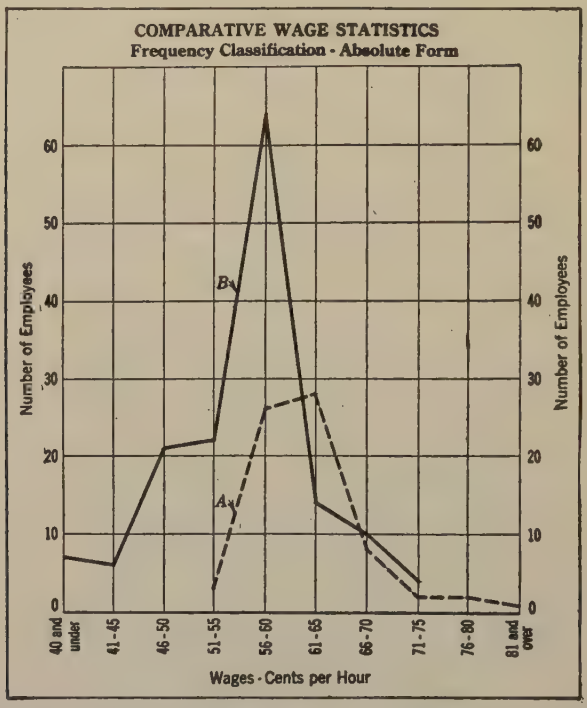


FIGURE 106.—Curve "A" Represents the Data, in Absolute Form, of Table 11, and Curve "B" the Absolute Figures of Table 10, both on page 398
Comparison of two frequency curves in absolute form.

for purposes of comparison, as illustrated by Figures 106 and 107, respectively.

The Cumulative Series.—Simple series are made cumulative when successive frequencies are added. In this case each class group contains not only its corresponding items but also the items contained in all of the class groups preceding it. The cumulation

may be made from the lowest value to the highest—in ascending order—or from the highest to lowest—in descending order. The former is called a "less than" cumulative series and is illustrated in columns "d" and "e" in Table 1, page 381, and in "A" of Figure 108 below. In this case the items in any group represent the number *less than* the measurement indicated by the class group or more

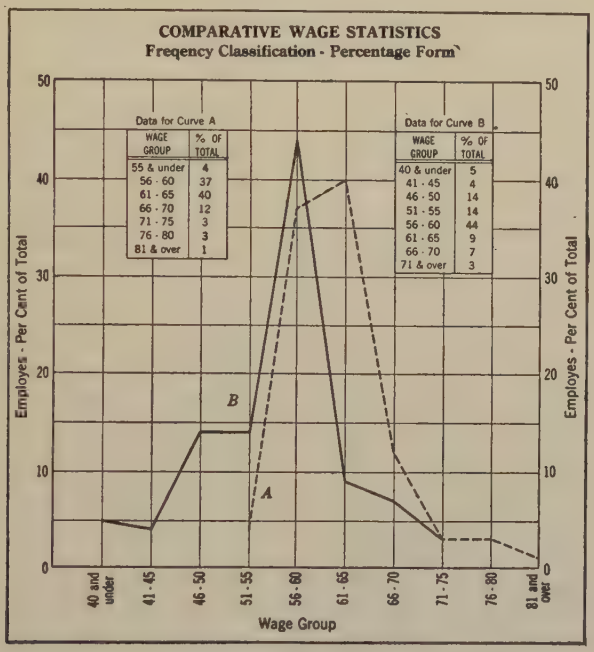

COMPARATIVE WAGE STATISTICS
Freqency Classification - Percentage Form
Data for Curve A
WAGE GROUP | % OF TOTAL
55 & under | 4
56 - 60 | 37
61 - 65 | 40
66 - 70 | 12
71 - 75 | 3
76 - 80 | 3
81 & over | 1
Data for Curve B
WAGE GROUP | % OF TOTAL
40 & under | 5
41 - 45 | 4
46 - 50 | 14
51 - 55 | 14
56 - 60 | 44
61 - 65 | 9
66 - 70 | 7
71 & over | 3
Employes - Per Cent of Total
Wage Group
B
A
40 and under
41 - 45
46 - 50
51 - 55
56 - 60
61 - 65
66 - 70
71 - 75
76 - 80
81 and over

FIGURE 107.—The Data of Figure 106 Expressed in Percentage Form
The two curves are much more readily compared than those of Figure 106 above.

strictly *less than* the upper class limit of the class group. Thus in Table 1, page 381, there are 136 items *less than* 67 or *less than* the upper class limit of this group. The latter is known as the "more than" cumulative series and is illustrated in columns "f" and "g" in Table 1, page 381, and in "B" of Figure 108 below. Here the items in any group represent the number *more than* the measurement indicated by the class group or *more than* the

lower limit of the class. In Table 1, on page 381, there are 430 items *more than* 67 or *more than* the lower limit of this group.

The original entries of the series may likewise be expressed in absolute numbers or percentages of the total as shown in columns "b" and "c" of Table 1 on page 381. The percentage form is required generally where different series are to be compared as illustrated by Figures 106 and 107, respectively. The cumula-

Hourly Rate of Pay—Men Factory Workers

Period of Service—2 Years

Wage Groups	No. of Employes	
	Absolute	% of Total
a	b	c
40 and under	7	5
41 – 45	6	4
46 – 50	21	14
51 – 55	22	15
56 – 60	64	43
61 – 65	14	9
66 – 70	10	7
71 and over	4	3
	148	100

Table 10

Period of Service—5 Years

Wage Groups	No. of Employes	Per Cent of Total
a	b	c
55 and under	3	4
56 – 60	26	37
61 – 65	28	40
66 – 70	8	12
71 – 75	2	3
76 – 80	2	3
81 and over	1	1

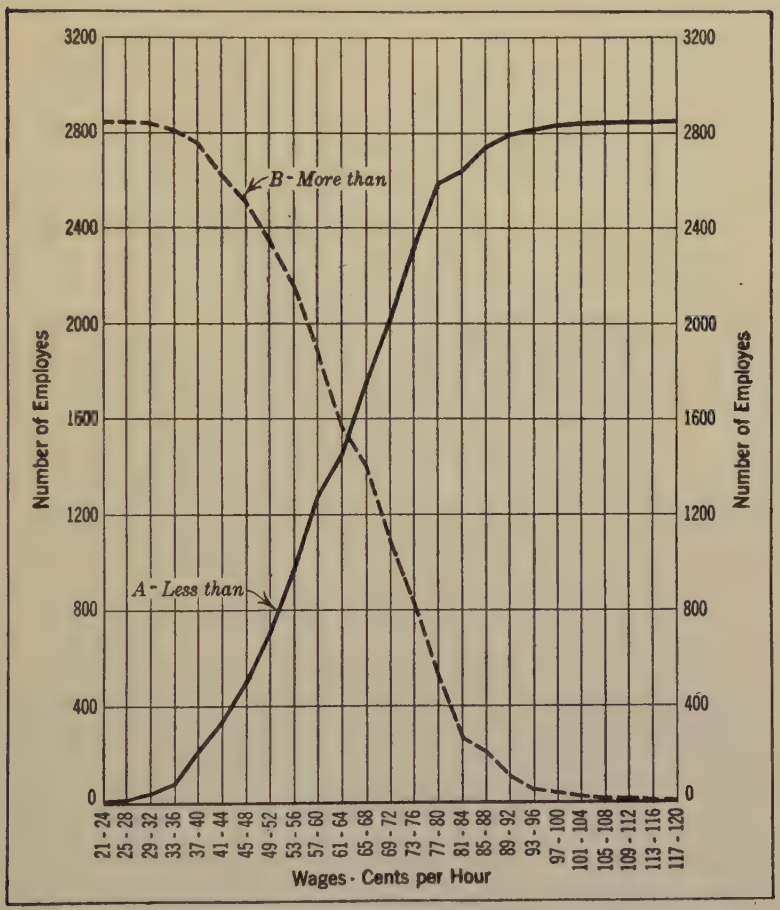
Table 11

tive form is serviceable in many ways. It is helpful in the computation of the median, quartiles, deciles, and so forth, indicated on page 418, and in other respects as noted in the next section.

Characteristics and Use of Frequency Analysis.—The chief characteristic of the frequency series is that generally the items tend to group about a value or measurement which is generally taken as the type, although there are exceptions as noted on pages 394 and 395. In practical business statistics frequency series are generally asymmetrical in character.

There are many cases in practical business statistics where

frequency classification will be employed in the process of analysis. While the general method of procedure outlined in this chapter will be followed in such cases, the details of the classification will be



(Data: National Metal Trades Association—N. Y. & N. J. Branch)

FIGURE 108.—Hourly Rates of Pay—Metal Trades Employes
Cumulative frequency curves or ogives.

developed from the standpoint of the practical aspects of each particular problem. The following will suggest the extent of the use of frequency analysis and are typical of its use in the field of internal facts:

PERSONNEL FACTS

1. Number of employes by:
   - Period of service
   - Age
   - Rates of pay
2. Average Rates of pay by:
   - Period of service groups
   - Age groups
3. Absences by:
   - Period of service
   - Age
   - Rates of pay
   - Length of absence
4. Labor losses by:
   - Period of service
   - Age
   - Rates of pay

FINANCIAL FACTS

1. Loans by:
   - Interest rates
   - Amounts
   - Time
2. Industrial securities by:
   - Dividend or interest rates
   - Amounts
3. Sales on stock exchange by:
   - Price
   - Size of sale in shares

OPERATIONS FACTS

1. Purchases by:
   - Price
   - Quantity
2. Sales by:
   - Selling price
   - Quantity in each sale
   - Size of article
3. Production facts:
   - Factory orders, shop deliveries, and unfilled orders by quantities
   - Shop deliveries by average cost groups
   - Time studies

From the frequency classification we can generally determine at once the most typical fact and broadly how the items of the series are grouped about it. By noting the most typical fact, the range, and the way the items are grouped generally in the series—whether they cluster closely about the type or are scattered widely from it, whether they are dispersed regularly or irregularly from the type, whether they are mostly above the type or below it—we can get a good general idea of the characteristics of the series. We can also tell by inspection the absolute and relative number of items in any class or group of classes and from the cumulative form we can tell the number above or below any point that is particularly significant, where ¼, ½, or ¾ of the items fall, and so forth.

While all this is very serviceable, it is not generally sufficient.

In most cases, the frequency classification has to be supplemented by computations of different kinds, such as averages, index numbers, ratios, and so forth. These computations are developed in subsequent chapters.

Conclusion.—The foregoing is but a brief presentation of frequency analysis. Particularly is the discussion of the frequency graph limited. For further discussion of the general subject and the frequency graph the student is referred to the references listed below. Mr. Karsten's treatment is particularly interesting in that it gives more elaborate consideration to the use of the logarithmic or rate-of-change scale in plotting frequency series.

References

Brinton, Willard C. Graphic Method for Presenting Facts, 1914, Ch. X.
Forsyth, C. H. Mathematical Analysis of Statistics, 1924, pp. 83-88.
Jerome, Harry. Statistical Method, 1924, pp. 36-42.
Karsten, Karl G. Charts and Graphs, 1923, Ch. XXVII, pp. 308-325 (Series), Ch. XXVIII, pp. 326-340 (Curves), Ch. XXIX, pp. 341-355 (Ogives).
Kelley, Truman L. Statistical Method, 1923, pp. 9-15 (Frequency Polygon).
King, Willford I. Elements of Statistical Method, 1912, Ch. XI.
Mills, Frederick C. Statistical Methods, 1924, Ch. III.
Rietz, H. L. Handbook of Mathematical Statistics, 1924, pp. 20-23.
Rugg, H. O. Statistical Methods Applied to Education, 1917, Chs. IV, VII (Graphs).
Secrist, Horace. An Introduction to Statistical Methods, 1917, pp. 144-153, 198-220 (Graphs).
West, C. J. Introduction to Mathematical Statistics, 1917, Ch. III.
Yule, G. U. An Introduction to the Theory of Statistics, 1919, Ch. VI.

CHAPTER XXXVI

AVERAGES—THE MEAN

The Meaning of the Average.—The *average* is one of the most extensively used computations in statistical work, being developed in practically every phase of personnel, financial, and operations statistics. We speak of the average rate of pay, average number of employes, average length of absences, and the average age; the average monthly income and expenses, average prices of stocks and bonds, average rates of interest, and average bank balances; and the average factory output, average sales, average deliveries, average costs, and so forth. Throughout all phases of the management of a business organization we speak of and think in terms of averages.

Most of the literature on the subject of averages has been devoted to the theoretical aspects of the different types and in most cases extensive attention has been given to fine points of distinction which have little interest to the practical business statistician. Without deprecating the development of methods in the slightest, emphasis will be placed in the present discussion on the practical application of these averages in business management. We do not generally require as fine a degree of accuracy in practical business statistics as is required in some of the other fields such as biology, anthropology, and astronomy. Hence, our chief concern is to determine the average which will best serve our purpose even if it is not so accurate as others. The use of the arithmetic average, for example, has been so general in business management that we must show where its use will lead to serious error before we can expect practical men to adopt other types.

Likewise, the general idea of an average calls merely for a summarizing expression but a specific purpose to be served may require that the summarizing expression possess certain specific characteristics, the presence of which is indispensable to this particular purpose. These variations in the specific characteristics

of the average are secured through the use of different methods of computing the summarizing expression. The different methods of computing the average produce single expressions which differ in detailed representativeness. For convenience, these different results are referred to as different types or kinds of averages rather than as averages arrived at by different methods.

In this connection, it may be noted that there is some confusion in the use of the terms *average, type* and *mean.* They are sometimes used synonymously and sometimes separate and distinct meanings are attached to them. The word *mean* in particular is used to designate the arithmetic average and again to represent the general concept of averages. In the present discussion, the term *mean* will be used when standing alone to designate the arithmetic average but it will be used synonymously with *average* when accompanied with the proper descriptive terms, as in the expression, the geometric mean. *Type* in this special sense will not be used at all.

Purpose of an Average.—In attempting to use the scientific method in practical business management we frequently find ourselves confronted with huge masses of data that have a significant bearing on the important problems with which we have to wrestle: but so numerous and complex are the facts that we find great difficulty in determining specifically and definitely just what that bearing is. In making comparisons and otherwise utilizing large groups of facts, the mind cannot handle masses of detail. They are burdensome and unwieldy. Suppose the problem were to compare rates of pay of the employes in two banks, A with 2,500 employes, and B with 3,000 employes. For us to attempt to determine this comparison by placing side by side 2,500 items representing the weekly rates of pay of the employes in Bank A and 3,000 items representing the weekly rates of pay of the employes in bank B, would be folly.

But this problem is typical of many that we meet every day in practical business management and the unwieldy character of these large groups of data necessitates that they be reduced to a single expression which gathers into itself the significant characteristics of the complex data. By means of this single expression

we can bring the effectiveness of these important groups of data to serve us in our efforts to solve practical problems.

Now, the average is the most generally satisfactory statistical device invented to date for determining this single summary expression. This device is not unfamiliar. It is a habit of mind to think in terms of averages. The expression "on the average" is a common one even among those who claim no technical knowledge of averages. To the casual questions, "What are you sales?" "What is your daily output?" "What is your overtime?" and so forth, the answer invariably runs, "on the average, so and so." The mind finding itself up against a mass of facts almost invariably resorts to the use of the average and since we are constantly confronted with large groups of data, we have very extensive need for the average in business management.

Types of Averages.—There are a number of averages which vary in method of computation and significance, as developed at some length on pages 131 *ff.* of Zizek's "Statistical Averages" (Translated by Persons). Our discussion will be confined to those particular averages which are used to a greater or less extent in practical business statistics.

Specifically we shall consider five different types of averages, distinguished by different methods of computation. These are as follows:

1. *Arithmetic Average or Mean*—Arrived at by dividing the sum of the items by their number.
2. *Median*—Arrived at by arranging the data in an array—*i. e.*, in order of size from lowest to highest or *vice versa*—and taking the middle value.
3. *Mode*—Arrived at broadly by taking the most frequent value in a group of representative data.
4. *Geometric Average or Mean*—Arrived at by multiplying the several items together and extracting a root of the product indicated by the number of items.
5. *Harmonic Average or Mean*—Arrived at by taking the reciprocal of the arithmetic average of the reciprocals of the items in the group.

It should be kept in mind that all of these averages are alike

in that they stand for detailed facts—that they are single expressions which embody in them to a greater or less degree the essential characteristics of the entire group of data which they may respectively represent. But they vary in the completeness with which they embody these essential characteristics and it is this feature which permits us to use different ones for different purposes.

The fundamental problem in the use of the average has to do with its representativeness; generally speaking, we want that particular average which is most representative from the point of view of the purpose to be served and which will consequently produce the most satisfactory results. The arithmetic average lays claim to representativeness because it is the result of leveling all the items in the group, arithmetically; the median, because, broadly, there are just as many items above as below it; the mode, because there are more actual items at this value than at any other; and the geometric, because it is the result of leveling all the items in the group, geometrically.

These different bases are significant in themselves and have an important bearing on the representativeness of the average aside from the difference in actual results from the different methods. In this connection the reader may find it interesting to note the requirements of a good type of average outlined on page 108 of Yule's "Introduction to the Theory of Statistics," although such tests are generally not sufficiently concrete to be of great value in practical work.

Furthermore, it should be noted that the different types of averages tell different stories and hence a given purpose may require the computation of more than one type.

Further Considerations.—Thus far we have considered the general aspects of the average. Our aim has been for the most part merely to lay the foundation for further development of the subject. This and the next three chapter will be devoted to a consideration of

1. The nature and method of determining each type of average.
2. The specific use of each type in the solution of practical business problems.

THE ARITHMETIC AVERAGE OR THE MEAN

The Nature and Computation of the Mean.—The specific definition which distinguishes the arithmetic mean from other types of averages is a descriptive one—viz., that the arithmetic average or mean is the result obtained by dividing the sum of a group of numerical data by the number of separate items included in the sum. It is a sort of leveling process. In effect, it decreases the larger numbers and increases the smaller ones and thus equalizes all of the numbers in the group.

This type of average is so familiar that its method of computation does not require elaborate treatment. The method is clearly implied in the definition given above. We simply take the sum of all the items in the group and divide this sum by the number of items. The process is simply illustrated in Table 1 below.

FINANCIAL STATISTICS.
Gross Revenues—By Months, 1923

Month	Gross Revenues
a	b
January	$ 41,215.75
February	39,105.40
March	42,300.00
April	42,943.25
May	43,175.35
June	43,560.60
July	43,725.80
August	43,904.75
September	44,008.50
October	44,250.90
November	44,385.30
December	44,425.20
Total for year	$517,000.80

TABLE 1.—To get the monthly average, the items in column "b" are totaled and divided by the number of months. Thus, $517,000.80 ÷ 12 = $43,083.40

We have here a series of 12 items. The sum of the items is indicated at the foot of the second column and the average is obtained by dividing the sum by 12. This simple example illus-

trates the fundamental method of computing the arithmetic average.

There are, however, special forms of making this computation which differ from the procedure detailed above. These special forms, however, are merely modifications or variations of the arithmetic process used—merely different methods of obtaining the same results—and in no way alter the fundamental method of computation nor the character of the resulting average. They are generally known as "short-cut" methods of computing the arithmetic average and are designed merely to facilitate the process of computation. Those chiefly used are (1) the weighted average method, (2) the assumed or guessed average method, and (3) the step method. The last two are discussed in Appendix E.

The Method of Weighting.—The short-cut method used most extensively is what is known as the method of weighting. In this process also we divide the sum of the items in the group by their number. It differs from the simple process noted in the previous section only in the manner of obtaining the sum of the items. Here, we obtain the sum of the items by multiplying each different value by the number of times it appears in the group and summating the products as shown in Table 2 below. The items of $100 and $105 occur 3 times, as shown in Table 3 below and are each multiplied by 3: $115 occurs 4 times and is multiplied by 4. An inspection of these two tables will reveal the true character of this method.

In Table 3 the prices for the 10 shares are listed separately and their total taken. In Table 2 each different price is multiplied by the number of shares to which it applies and the products totaled. The two sums thus obtained are equal.

A comparison of these two tables shows that there is nothing new, fundamentally, in this method of determining the arithmetic average. It is simply a short method of performing the arithmetic of the process. Multiplication is a short method of addition and we have utilized it in Table 2 instead of addition. This is the only difference between the weighted method and the simple method noted in the preceding section.

The numbers by which the different values are multiplied are

called "weights" and represent the relative importance given to each item in the total. These *weights* may be arrived at by a process of grouping as in Table 2 or they may be arrived at in ways entirely different without relation to the series. The method of selecting the weights, however, does not affect the process of computing the average. In any event the average is the result obtained by dividing the number of items or the sum of the "weights" into the sum of the products obtained by multiplying the items by their respective weights.

In this connection, it should be emphasized that questions arising in connection with the selection of the weights are not a part

SHARES OF STOCK—PRICE

Weighted Process

Price per Share	No. of Shares	Product
$100	3	300
105	3	315
115	4	460
Total	10	$1,075

TABLE 2.—Arithmetic Average—$1,075 ÷ 10 = $107.50 average price per share

Simple Process

$ 100
100
100
105
105
105
115
115
115
115
$1,075

TABLE 3.—Arithmetic Average—$1,075 ÷ 10 = $107.50

of the process of computing the average. The computation of the average begins only after the weights have been selected. Weighting the items in a group simply consists of determining their relative importance and the process of computation will be the same regardless of what this relative importance may be. The question of weighting in statistical analysis is a subject by itself and is not confined exclusively to the computation of the arithmetic average. We are obliged to determine weights in many other cases and the question of weighting should be dissociated from the method of computing the arithmetic average.

Weighting and the Frequency Series.—The weighted procedure outlined above is the regular method of computing the arithmetic average in a frequency series. Table 2 above is a sim-

ple form of frequency grouping. Table 4 below, however, is more typical of the frequency series and will serve to illustrate the use of weighting in such cases.

WAGE STATISTICS.
Base Rates of Pay—Weekly Basis.

Wage Groups	Mid-Points of Groups	Number of Employes (Frequencies)	Products
a	b	c	d
$10 under $15	12.00	10	120.00
15 " 20	17.00	15	255.00
20 " 25	22.00	50	1,100.00
25 " 30	27.00	12	324.00
30 " 35	32.00	8	256.00
35 and over	37.00	5	185.00
Total		100	2,240.00

TABLE 4.—To obtain the average (arithmetic) rate of pay, the mid-points (see page 386), of the wage groups are multiplied respectively by the number of employes (weights) and the sum of these products is divided by the total number of employes (sum of the weights).

The Use of the Arithmetic Average.—Most of the cases in practical business statistics which require an average necessitate the use of the arithmetic average. Thus the average monthly revenues and expenses, the average daily sales, the average weekly output, the average number of employes, the average plant in service, and a multitude of similar computations require generally the arithmetic average. These are mostly cases of such nature as to require an average from which by simple operation the total figures can be derived. This is possible only with the arithmetic average and hence the other types cannot be used at all.

Furthermore, in the computation of the so-called *isolated averages* the arithmetic average is the only one which can be computed—*i. e.*, when we have only the sum of the items and their number. This is true generally in such cases as computing average daily revenues or expenses, average rates of pay, and so forth, where the records show the total figures only. Other typi-

cal examples are per capita consumption of various commodities such as meats, sugar, gasoline, tobacco; average number of letters, periodicals, money orders, and so forth, per individual; average floor space per occupant in an office; average expenditures, receipts, or debt of a country per capita; and so forth.

This type of average is discussed at some length in Zizek's "Statistical Averages" (Translation by Persons), beginning at page 27.

In fact, the arithmetic average is used most extensively in practical business statistics and it cannot be too strongly emphasized that aside from those problems which require specific types of averages as noted in Chapters XXXVII, XXXVIII, and XXXIX the arithmetic average should be used generally. It is more adapted to problems of business management and when the range is associated with it we get a very serviceable representation of the group of data.

It must be noted also that we cannot entirely disregard those whom we are serving. Business executives are generally more familiar with the arithmetic average and in many cases will insist that it be used. Nevertheless, lack of familiarity is not a final reason for not using a particular type of average. If it is the proper average to use we should persistently though tactfully urge its use until it has become familiar and its service is fully appreciated.

Care should be taken in the application of the arithmetic average. In practical work there is a tendency to use short cuts which may lead to serious error. For example, it is a favorite practice to find the monthly average—say, of employes or capital investment, or bank balances—by dividing the sum of the figures at the beginning and end of the month by 2. This may be satisfactory for practical purposes but the possibility of error is evident. This computation is based on the assumption that the behavior of the data has been somewhat regular throughout the month—that the change indicated by the difference between the two figures has been gradual. This may or may not be true and the detail should be carefully inspected as a rough check on the method. The same will apply to the method of obtaining averages for the year or other time period by averaging the figures

at the beginning and end. We get a more accurate yearly average theoretically by averaging the monthly figures than by taking only the two and we should improve the results still more by using still greater detail.

The aim here is not so much to argue against such practices as to urge caution in their use. They are subject to the same criticism as the method of averaging a series by averaging its extremes—the details in the time period are entirely disregarded in the one case and those in the body of the series in the other.

(For list of references see page 435.)

CHAPTER XXXVII

AVERAGES—THE MEDIAN

The Nature of the Median.—It was stated on page 404 that a second method of determining a figure to represent a group of data was to arrange the figures in an array and select the middle value. This representative figure or average is called the *median*. Its representativeness, broadly speaking, is based on the fact that there are just as many items in the series which fall below it in value as there are items which fall above it. In other words it divides the number of items in the series or array into two equal parts. Generally, it is the value in a statistical array on either side of which half of the items are located. It is simply the balancing point of a series on the basis of the *number* of items just as the arithmetic average is the balancing point on the basis of the *value* of the items.

It is important to note that the median may be fictitious and not one of the measurements of the series. This is illustrated by Tables 1 and 2 below. In the one series, the number of items is odd and the median is a real item, 23. In the other, the number of items is even and the value of the median is 22½, half-way between 22 and 23, or the mean of these two.

How the Median is Computed.—The fundamental method of computing the median is evident, comprising only the location of the middle value in the array. Because of the difference in the complexity of the two processes, we shall consider the method of computation separately for, (1) a series in which each measuremen or number occurs only once—*i. e.*, in which the frequencies are unity for each value, and (2) a series in which the measurements or values occur more than once—*i. e.*, in which the frequencies in the several classes are two or more. The former is illustrated by Table 1 below, and the latter by Table 3 below.

It is important to note also that the computation of the median is affected by the character of the series—*i. e.*, whether the

data are continuous or discontinuous as distinguished on page 380. The importance of this point will be noted later in the discussion.

It should be emphasized here that no attempt should be made generally to typify a series when the number of items is evidently too small to be representative. Likewise, an inspection should be made of the series to determine whether it lends itself to representation by the median. We find many series in practical work which cannot properly be thus represented.

The Simple Series.—This type of series has but a single frequency in each class. It is illustrated by Table 1 below. There are three principal steps in the location of the median in a series of this type:

1. *Array the Data*—This is the process of arranging the data in order of size, either ascending or descending.
2. *Determine the Median Term*—This is accomplished by the formula $\frac{n+1}{2}$ in which n represents the number of items in the series.
3. *Locate the Median Value*—The value which corresponds to the middle term is determined by simply counting in from either end of the series. If the number of items in the series is *odd* the median value will be an actual item: if even, it will fall between the two middle-most items. By established practice, in the latter case we take the mean of the two middle-most items as the median, (the value half-way between them) although in the continuous series it may actually fall at any other point between them. This will be true even though there are gaps in the series: if, for example, 12 and 13 are added to the series of Table 2 below and the items, 21 and 22 omitted, the median value will be the average of 20 and 23 or 21.5.

The process is illustrated by Tables 1 and 2 below.

Quartiles, deciles, and percentiles may be determined in the same general way in which the median is located. Thus $\frac{n+1}{4}$ and $\frac{3(n+1)}{4}$ are the formulas respectively for the lower and upper or first and third quartiles. Applying this formula to the data of Table 1, Q_1, the first quartile, is 18 and Q_3, the third quartile, is 28. Likewise, $\frac{n+1}{10}$, $\frac{2(n+1)}{10}$, $\frac{3(n+1)}{10}$, etc., are the formulas

(Number of Items, Odd)
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32

TABLE 1

Number of Items = 19.
Substituting in formula, $\frac{n+1}{2}$,
we have $\frac{19+1}{2}$ or $\frac{20}{2}$ or 10. Hence the 10th term is the median term. Counting up from the lowest item or down from the highest, the 10th term is 23, the median value.

(Number of Items, Even)
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31

TABLE 2

Number of items = 18.
Substituting in $\frac{n+1}{2}$ we have
$\frac{18+1}{2}$ or $\frac{19}{2}$ or 9½. Hence, the 9½th term is the median term. Counting up from the lowest item or down from the highest, we find the 9½th term to be between 22 and 23. By conventional practice we take the value half-way between these values—*i.e.*, 22.5 is the median value.

for computing deciles. Applying these formulas to the data of Table 1, the first, second and third deciles are respectively 15, 17, and 19.

When the number of items in a series of discontinuous data is even we are confronted with the difficulty of obtaining a median which is fictitious in every sense of the word. The very character of the data makes it impossible for a half of an item to exist in fact. In such case the median should be used only with its true nature clearly in mind.

In this connection, the general idea of the median noted on page 412 requires qualification. It is true that the median is the middle value of the series but when we add that it divides the series into two equal parts the statement is not quite correct. This is particularly so in the discontinuous type of data as pointed out by Yule on page 116 of his "An Introduction to the Theory of Statistics." In Table 4 below, for example, which contains an odd

number of items, the median value is 6 but there are approximately 49% of the items below this and approximately 35% above. Similar illustrations might be noted from different types of data in the practical field.

The Grouped Series.—The grouped series is illustrated by Tables 3 and 4 below. The method of procedure in determining the median in this type of series is broadly the same as that followed in the simple series. The fact that the items have been grouped into classes offers a complication in determining the median value after the median term has been found. It should be kept in mind in working with the grouped series that the median tends to divide the *frequencies* or *items* into two equal parts and not the *measurements*. Specifically, the steps in this process are as follows:

1. *Array the Measurements*—Column "a" in Tables 3 and 4.
2. *Locate the Median Term*—This is accomplished as in the simple series by applying the formula $\frac{n+1.}{2}$[1]
3. *Locate the Median Group*—The only problem involved in this step is to determine the group which contains the middle term. This is done by the simple process of counting from the lower or upper end of the series until the group is reached which contains the middle term. The preferable method of procedure is to cumulate the series in which case the median group is readily determined by inspection. Column "c," Table 4.
4. *Locate the Median Value*—The method of determining the median value within the median group is based on the assumption that the items in the group are equally distributed in order of size throughout the group—*i. e.*, that they regularly increase in value from the lower limit of the group to the upper limit. On the basis of this assumption we simply take such portion of the class interval of the median group as the number of items required from this group to reach the median term is of the total in the group. Specifically, the procedure is as follows as shown in Table 3 below:
   (1) Summate the items contained in all the groups below or above the median group.

[1] It seems clear that this is the proper method of determining the middle term in spite of the fact that the question is sometimes raised as to whether $\frac{n}{2}$ is not preferable.

(2) Subtract the number determined in (1) from the median term. The result represents the number of items that must be taken from the median group to reach the median.

(3) Take such part of the class interval of the median group as the result of (2) is of the total number of items in the median group—*i. e.*, divide the result of (2) by the total number of items in the group and multiply the quotient by the class interval.

(4) Add the result of (3) to the lower limit of the class or subtract it from the upper limit as the case may be. The result is the median value.

WAGE STATISTICS.

Base Rates of Pay—Weekly Basis.

Wage Groups	Number of Employes (Frequencies)		Products
	Consecutive	Cumulative	
a	b	c	d
$10 under $15	10	10	120.00
15 " 20	15	25	255.00
20 " 25	50	75	1,100.00
25 " 30	12	87	324.00
30 " 35	8	95	256.00
35 and over	5	100	185.00
Total	100		2,240.00

TABLE 3.—In determining the median value of this series of wage data the three steps are as follows:

(1) Median term is located by dividing the number of terms plus one by two—*i.e.*, $\frac{n+1}{2} = \frac{100+1}{2} = \frac{101}{2} = 50\frac{1}{2}$.

(2) Median group is that group which contains the $50\frac{1}{2}$th term. By inspecting the cumulative entries in column "c" it is apparent that this is the group 20–25.

(3) Median value is approximated by interpolating within the median group on the assumption of an equal distribution within the class. Hence, following the procedure outlined above.

$50\frac{1}{2} - 25$ (number of items below median group) $= 25\frac{1}{2}$, number of items required in the median group.

$\frac{25\frac{1}{2}}{50}$ of 5 = 2.55 distance the median value is from lower limit of the group.

∴ 20 + 2.55 = $22.55, the median value or wage.

In a grouped series of the discontinuous type, interpolation may produce a fictitious value. In such case, if the details are at hand it is preferable to choose an actual value from an examination of the detail. Otherwise the computed value may be used but in such event it should be used only as a *fictitious* value.

NUMBER OF SALES BY SIZE OF ARTICLE.
1920

Size of Article	Number of Sales	Cumulative Sales
a	b	c
1	65	65
2	120	185
3	237	422
4	379	801
5	222	1,023
6	336	1,359
7	228	1,587
8	165	1,752
9	134	1,886
10	102	1,988
11	63	2,051
12	25	2,076
13	19	2,095

(Data from Figure 103)

TABLE 4.—Median term is (2095 + 1) ÷ 2, or 1048, and the median group is 6. Since the group is discontinuous and the class internal is equal to the unit of measurement, 6 is also the median value.

Note that approximately 49% of the items fall below this value and approximately 35% above it, and that the median, therefore, does not divide the series into two equal parts.

The same method of procedure will be followed in determining quartiles, deciles, and other such measurements in the grouped series as has been outlined above for the median.

Graphical Interpolation.—It is an advantage of the median that it can be determined roughly from the cumulative form of the frequency curve. The steps to be followed are as follows:

1. Locate the median term on the vertical scale by inspection. This is 3097 in Figure 109 below.
2. From the point determined in No. 1, draw a line horizontally to the right until it intersects the ogive, line "M" in Figure 109 below.
3. From the point of intersection drop a perpendicular to the

horizontal base. The point where this perpendicular strikes the base is the median value. Point P_2 in Figure 109 below is the median.

When the curve is smoothed to approach the distribution of the entire "universe," the median so located approaches the median of the "universe." Comparing this with the method of computa-

HOURLY RATES OF PAY—MACHINE SHOP EMPLOYES.
October, 1924.

Cents per Hour	Number of Employes
25 – 28	16
29 – 32	44
33 – 36	86
37 – 40	240
41 – 44	316
45 – 48	778
49 – 52	1,398
53 – 56	2,135
57 – 60	3,145
61 – 64	3,497
65 – 68	4,322
69 – 72	5,102
73 – 76	5,586
77 – 80	5,924
81 – 84	5,981
85 – 88	6,080
89 – 92	6,120
93 – 96	6,161
97 – 100	6,168
100 – 110	6,193

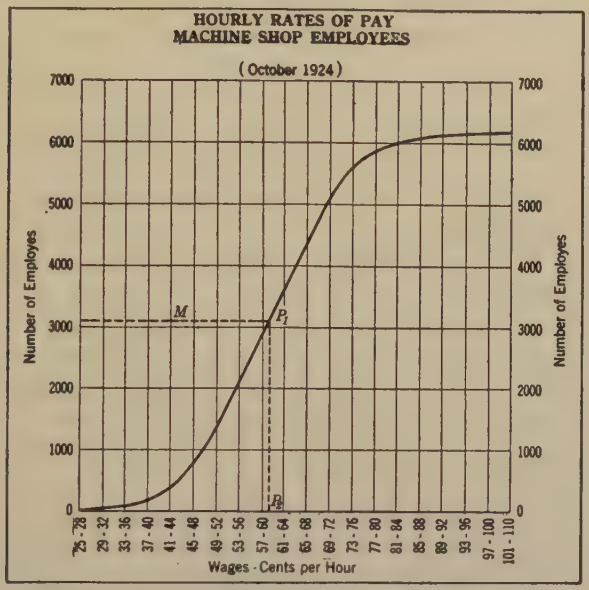
(Data: Metal Trades Association—N. Y. & N. J. Branch)

TABLE 5.—Number of Employes Cumulated (Data for Figure 109)

tion, Professor Yule states on page 118 of his "Introduction to the Theory of Statistics," that if the curve is flat there will be no appreciable difference between the results obtained by the two separate methods. If the curve is steep, the graphical value is preferable, "as it does not involve the crude assumption that the frequency is uniformly distributed over the interval in which the median lies."

Quartiles, deciles, and so forth, may be located in similar manner.

The Use of the Median.—The median is taken as an average simply because, generally speaking, it is the middle term—that is, there are just as many items below it as there are above it. Its representativeness depends greatly on the type of the series and its computation should be made only after an inspection of the series justifies it. It has a very limited use in practical work.



(Data: National Metal Trades Association—N. Y. & N. J. Branch)

FIGURE 109.—A Less-Than Cumulative Curve

The dotted line drawn from the median term 3097 on the vertical scale intersects the curve at the point P_1. A perpendicular dropped from point P_1 indicates the media value at the point P_2 on the horizontal scale. The quartiles may be located in a similar manner.

When taken with the quartiles it is resorted to occasionally in describing the middle half of the items in an array.

It is particularly serviceable in those rare cases where the data, such as may be found in measuring the ability of employes or in endurance tests, advertising campaigns, and similar cases, cannot be expressed in definite measurements but are expressed in order of rank.

(For list of references see page 435.)

CHAPTER XXXVIII

AVERAGES—THE MODE

General Conception.—The word "mode" is not unfamiliar. In general use it signifies the customary, the usual, the fashionable, the thing that is most generally done. In the technical use of the term in statistical methods, it carries the same general significance. Broadly speaking, it is used to designate the most frequent value in a series of statistical data, or that value which occurs the greatest number of times in the group. As Professor Yule states, the mode is the "value which is most frequent or typical, the value which is in fact the fashion (la mode)." In Table 1 below, for example, 24 is the mode since it occurs a greater number of times than any other measurement in the series. This concept, however, is only generally true. The correctness of its representation varies with what are known as the "true" and the "approximate" mode and with the representativeness or "randomness" of the group of data.

Waist Measurements of 2,139 Women, Pratt Institute

Waist in Inches	No. of Students
20	18
22	392
24	933
26	564
28	159
30	41
32	19
34	10
36	1
38	2
Total	2,139

(G. E. F. Saum in *Management and Administration*)

TABLE 1

The True Mode.—Invariably in practical work the refinement to which measures are taken is determined arbitrarily and the widths of the classes into which data may be grouped for purposes of analysis are also arbitrarily fixed. It is obvious that these arbitrary decisions will materially affect the relative number of frequencies in each class. By arbitrarily changing the class-limits we can have different values in the series occurring most frequently. That is, we can shift the modal value to a certain extent simply by changing the arbitrary features of the classification.

This condition is overcome only by having an adequate number

of items in the series, organized into classes of sufficiently small class-interval to give a so-called ideal distribution—*i. e.*, a distribution in which the class-intervals are small and the distribution nevertheless smooth. In other words, in practical work, generally if we have small class-intervals we get an irregular distribution and if we enlarge the class-interval to smooth the distribution errors enter in, due to the arbitrary selection of the class limits. To approach the "ideal" distribution we must have both very small class-intervals and a smooth distribution. This requires a very large number of data.

The actual attainment of such a series is obviously impossible in practical work. Consequently, the practical method of procedure is to determine the approximate value of the mode by fitting "an ideal frequency curve of given equation to the actual figures." The maximum value in this curve or the largest ordinate will be the "true" or "theoretical" mode. This, according to Professor Yule, is "the only strictly satisfactory method" of determining the mode.

Now, it is clear that the definition of the mode given in the section above will not be correct for this conception of the mode. The "true" mode will not necessarily coincide with the value in the series which is most frequent. When we have the "true" mode in mind such a definition will not hold. This will be apparent from the following definition of the "true" mode taken from page 120 of Yule's "Introduction to the Theory of Statistics": "The mode is the value of the variable corresponding to the maximum of the ideal frequency curve which gives the closest possible fit to the actual distribution." The fundamental conception of the mode holds true here but it is the most frequent value of the ideal distribution of which the actual group of data is a sample.

The application of this method is too involved for the purposes this book is designed to serve. Those who are interested will find the method fully treated in the references listed on page 435. For our purpose, we must resort to another method which, though producing less accurate results, is nevertheless sufficiently accurate for practical purposes. The modal value determined by this less accurate method will be called the "approximate" as distinguished from the "true" mode.

The Approximate Mode.—In practice we generally undertake to determine the mode of the actual data, regardless of the ideal distribution. In this sense the mode is, broadly, the most frequent value in the group of data actually before us regardless of how well this series corresponds to the ideal series. This "approximate" mode is sometimes referred to also as the "inspection" mode. In common practice the term "mode" generally signifies this value and throughout these discussions it will be used in this sense unless otherwise specified.

This conception will incorporate the idea that the group of data must be large enough to approach a representative sample and that the data must be distributed about the modal value as a central tendency. For purposes of these discussions, therefore, the mode will be defined as the value or values in a series of data—sufficiently large in number to be representative and arranged in order of size—which occur substantially more frequently than the general run of values in the group and around which the values in the series are grouped in more or less regular fashion, the frequency of their appearance shading off gradually on either side. In case more than one mode exists the series is said to be bi-modal, tri-modal, and so forth depending on the number of modal values present.

Selecting the Classes in Which to Group the Data.—The modal group and consequently the modal value may be materially affected by the arbitrary decisions entering into the selection of the classes into which the data are grouped. Generally, in practical work we find it necessary to enlarge the classes so as to smooth the distribution and reveal the point or points of concentration. In doing this, any one of a number of classifications may be set up and the value of the mode may be greatly affected by the one actually selected. Consequently, great care should be taken properly to determine these classes.

This point is considered at some length in Bowley's "Elements of Statistics," (2nd Edition) page 119 *ff.*, and also in Zizek's "Statistical Averages" (Translation by Persons), page 234 *ff.*

How the Mode is Determined.—As in the case of the median, the mode will be best computed by arranging the data in an array.

It is important here as in previous cases that the classes into which the data are grouped be of equal width. The determination of the mode involves three steps—(1) arraying the data, (2) locating the modal group, and (3) locating the modal value within the group. The method of arraying statistical data is discussed on page 384.

The Modal Group.—Generally, the modal group is readily determined from the array, by inspection. By looking over the series of frequencies we locate the largest group of items and the group value corresponding to this is the modal group. Table 2 below is illustrative: the largest number of frequencies is 50 and its corresponding group value, 20-25, is the modal group.

It is necessary, of course, that the frequency series be reasonably smooth before we undertake to locate the modal group. Or, perhaps, we should say that it should be sufficiently smooth properly to reveal the modal group. If the modal group is not clearly apparent, it is necessary usually to re-group the data. The most common method of smoothing irregularities is to increase the class interval or to widen the classes into which the data are grouped as discussed on pages 386 to 388.

WAGE STATISTICS.

Base Rates of Pay—Weekly Basis.

Wage Groups	Mid-Points of Groups	Number of Employes (Frequencies)	Products
a	b	c	d
$10 under $15	12.00	10	120.00
15 " 20	17.00	15	255.00
20 " 25	22.00	50	1,100.00
25 " 30	27.00	12	324.00
30 " 35	32.00	8	256.00
35 and over	37.00	5	185.00
Total		100	$2,240.00

TABLE 2.—The arithmetic average = $2,240.00 ÷ 100, or $22.40. The mid-points in the groups are used in accordance with the discussion on page 386.

The Modal Value.—The problem of locating the modal value will vary with *discontinuous* and *continuous* series. Special attention is given to discontinuous series on page 380. The discussion at this point will be confined to the continuous series.

The modal value within the group is determined in two principal ways. We may take the mid-point of the modal group in which case the determination of the modal value is simple. Or, we may take into account the comparative distribution of the data in other classes above and below the group.

In the application of the latter method we may include the data in all the classes above and below the modal group or we may use only a part of them. In practice, we usually take into account only the first one or two groups on either side of the modal group. More than these may be advisable when the series is very irregular but not if the series is multi-modal. The number of classes to use can be determined only by inspection of the particular series.

In this case, we assume that the numbers below the modal class tend to pull the modal value toward the lower limit of the modal group and those above tend to pull it in the opposite direction. On the assumption, therefore, that the data in the modal group are equally distributed over the class, we divide the class-interval of the modal group in inverse proportion to the number of frequencies in the classes selected above and below the modal class and the dividing point is the modal value. The process consists of two steps:

1. Take such part of the class-interval of the modal group as the number of frequencies selected below the modal group is of the sum of those selected below and those selected above.
2. Subtract the resultant figure from the upper limit of the modal class.

Or,

1. Take such part of the class-interval of the modal group as the number of frequencies selected above the modal group is of the sum of those selected above and those selected below.
2. Add the resultant figure to the lower limit of the modal group.

These two processes will be illustrated by the data of Table 2 above. By inspection, 50 is the largest number of frequencies in any class and consequently 20-25 is the modal group. Using only a single class below and above the modal group—*i.e.*, classes 15-

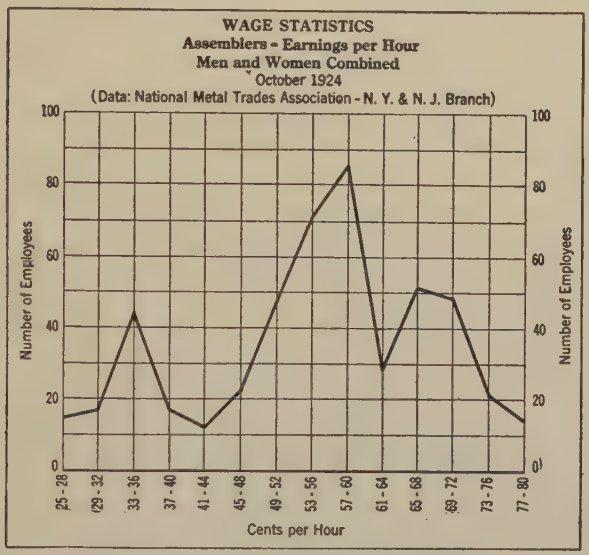


FIGURE 110.—A Tri-Modal Distribution
The principal modal group is 57-60, with sub-modes above and below as indicated.

20 with a frequency of 15, and 25-30 with a frequency of 12—the sum of the frequencies in the two classes is 27, and the procedure is as follows:

First Process—In this case, we take 15/27 of 5, the class-interval, and deduct it from 25, the upper limit of the modal class.

Hence, 15/27 of 5 = 2.78
and $25 - 2.78 = 22.22$, the modal value

Second Process—In this case, we take 12/27 of 5, the class-interval, and add the result to the lower limit of the modal class.

Hence, 12/27 of 5 = 2.22
and $20 + 2.22 = 22.22$, the modal value

As noted above, we may include more of the classes above and below the modal group and since the result will be materially

affected by this selection it is important that it be made properly. Generally speaking, unless the series is very irregular the use of the single class on either side is satisfactory.

If we wish to take the mid-point of the modal group, the modal value will be, roughly, $22.00.

Finally, in moderately asymmetrical series the modal value may be located directly from the arithmetic average and median with-

WAGE STATISTICS.
Assemblers—Earnings Per Hour.
Men and Women—October, 1924

Cents per Hour	No. of Employes
25 – 28	15
29 – 32	17
33 – 36	44
37 – 40	17
41 – 44	12
45 – 48	22
49 – 52	47
53 – 56	71
57 – 60	85
61 – 64	28
65 – 68	51
69 – 72	48
73 – 76	21
77 – 80	14

(Data: National Metal Trades Association—N. Y. & N. J. Branch)

TABLE 3.—A tri-modal distribution with concentration points indicated at 33-36, 57-60 and (65-68 and 69-72).

out attempting to locate the modal group. This method is based on the relationship existing between the three types in the moderately asymmetrical series, that the difference between the median and the mode is about 2/3 of the difference between the arithmetic average and the mode. Hence, three times the difference between the median and the arithmetic average equals the difference between the arithmetic average and the mode, which added to or subtracted from the arithmetic average gives the modal value.

Series with More than One Mode.—It was noted on page 422 that a series of data may have more than one mode—*i.e.*, it may

be bi-modal, tri-modal, or whatnot, depending on the number of modal groups. Table 3 and Figure 110 above, represent a typical tri-modal distribution. Generally, in practical work one of the modes is most prominent in the group, the others being relatively subdued and hence sometimes called sub-modes or secondary modes. They may, however, be of equal prominence.

The presence of more than one mode is generally an indication that the series is made up of two or more distinct groups of homogeneous data. Thus a series of wage statistics containing rates of pay for both men and women or for separate occupational classes may reveal more than one modal point. Likewise, series showing volume of sales which include sales of high-priced and low-priced articles, statistics on hours of work for office and factory workers, and so forth, will generally contain more than one mode. In such case it is well to separate the data into the separate groups, if it is feasible to do so, in which event each group will generally be uni-modal.

The Location of Mode in Graphs.—The value of the mode may be readily determined from an inspection of the frequency curve. In such a curve, assuming the measurements to be plotted horizontally and the frequencies vertically, it is the value which corresponds to the maximum ordinate. This will generally be only a rough measurement but it is usually satisfactory for practical purposes.

It may also be roughly located from the cumulative frequency curve or ogive. In this case it is the point at which the greatest number is added and hence at which the curve is the steepest. This point may be determined practically by passing the edge of a ruler along tangent to the curve and at the point where the ruler is most nearly vertical, or, the point where it ceases to turn in the one direction and begins to turn in the other, the modal value is registered. The value on the horizontal scale corresponding to this point is the modal value.

The Mode in Discontinuous Series.—Generally speaking, we do not have the difficulty in using the mode with a discontinuous series that we experience in other types of averages since the mode is always an actual measurement. However, when we

have a grouped discontinuous series, the problem of locating the modal value within the group becomes a difficult one. As a matter of fact such a refinement in connection with the discrete series should be undertaken only after the most careful inspection of the detailed data which make up the series. In actual practice, in classifying strictly discrete data we make the class-interval the same size as the measurement imposed by the character of the data and thus avoid the necessity of interpolating.

The Nature and Characteristics of the Mode.—The mode is taken to represent the group of data primarily because it is the *most probable value in the series.* When the data are arrayed, the items in the series are grouped around the mode, the number of frequencies being greater in the classes near the mode on either side and gradually falling off as the group values recede from the modal group. The probability that an item chosen at random will fall in the modal group is greater than that it will fall in any other group. This is the fundamental characteristic of the mode. It should be noted also that the mode depends on the size of items only at a certain point.

The Use of the Mode.—The use of the mode is likewise limited in comparison with the arithmetic average but nevertheless there are cases in which it is the only average that will serve the purpose. This is particularly true in cases which require that the average be an actual item or those in which the chief interest is in the "normal" or "typical" fact. This is frequently illustrated by referring to the method whereby the ready-made clothier should determine the size of suits to manufacture. It would indeed be a strange and profitless procedure for him to take the measurement of a number of men and combine these measurements by the arithmetic process to get the measurement of his suits. Rather should he take these measures and distribute them in a frequency series to determine the proportional number of each size similarly to the distribution of waist measurements shown in Table 1 page 420.

Furthermore, in developing a rate of pay for the business community by collecting and combining the rates paid by other companies, we may combine the rates of the several concerns

arithmetically by weighting them by the number of employes. But this will generally result in a fictitious rate which no one is paying and to say that it is the rate being paid generally in the community is ridiculous on its face. The proper average to use in this case is the mode, which represents the rate being paid by most concerns or to most employes according to the basis of the classification. There are many cases of this kind in practical work and they all definitely require the modal average.

Again, we may wish to determine the average annual income reported by citizens of New York State as given in the Federal income tax returns in the year 1921.. In this event if we want an average which is simply a condensed representation of the total, the arithmetic average will serve; but if we are interested in the welfare of the citizens of the state we should require the mode.

Similarly, if our interest in rates of pay of the employes of the company is from the financial or expense point of view we shall usually want an average which when multiplied by the number of employes will result in the total payroll—we want the arithmetic average; but if we are interested from the point of view of the employes, the mode will be much preferable.

In averaging groups of data of this general type and for certain purposes the mode is certainly the preferable average, if not the only one that can be properly employed, and in such cases it should be used.

(For references see page 435)

CHAPTER XXXIX

GEOMETRIC AND HARMONIC AVERAGES

Other Types of Averages.—The arithmetic average, the median, and the mode developed in the chapters immediately preceding, are the most commonly used types of averages. There are two other types, however, which are required in certain cases and which will be given brief consideration at this point. These are (1) the geometric average and (2) the harmonic average.

This chapter presents a somewhat concrete description of each of these types and points out the method of computing them.

The Definition and Nature of the Geometric Average.—Letting "n" represent the number of items in the group, the geometric average may be descriptively defined as the nth root of the product of the several items in the group. Like the arithmetic average, it may be considered a leveling process. The leveling process in this case, however, is of a different kind. Here we level the product of the items in the group of data while in the arithmetic average we level their sum. In this process we determine a figure which, raised to a power indicated by the number of the items in the group, will give the continued product of the several items in the group: in the arithmetic average we determine a figure which, multiplied by the number of items in the group, will give the sum of those items.

Likewise, in the process of computing the geometric average, the larger numbers in the group of data are reduced and the smaller ones increased until the differences are all leveled out; but the method of doing this is different from that followed in the arithmetic average. Instead of the absolute values as in the arithmetic average, we are in this case interested in the relative values of the individual items.

Computation of the Geometric Average.—The definition given above suggests the method of computing the geometric

average. In computing the arithmetic average, we add the several items and divide by their number. In computing the geometric average we multiply the several items together and extract a root indicated by the number of items in the group.

As an illustration let us take the group of abstract numbers, 8, 7, 5, 4, 2, 3. The arithmetic average of this group is computed as follows:

$$\text{Arithmetic average} = \frac{8+7+5+4+2+3}{6}$$
$$= \frac{29}{6}$$
$$= 4.83 -$$

And, the geometric average as follows:

$$\text{Geometric average} = \sqrt[6]{8 \times 7 \times 5 \times 4 \times 2 \times 3}$$
$$= \sqrt[6]{6720}$$
$$= 4.34 -$$

The method of procedure in computing the geometric average is simple. The actual carrying out of the process in practical work, however, is rather difficult. It is obvious that the computation, if followed through as suggested by the illustration, would be very burdensome where large numbers are involved. This difficulty is partially overcome by the use of the principles of logarithms noted on pages 503 to 505 for obtaining the products and roots of numbers. Specifically, the procedure is as follows:

1. By means of a table of logarithms, find the logarithm of each item in the series.
2. Summate these logarithms; the total thus obtained is the logarithm of the product of the several items.
3. Divide the total obtained in No. 2 by the number of items; the quotient is the logarithm of the nth root of the product of the items of the series.
4. By means of a table of logarithms, find the number which corresponds to the logarithm resulting as the quotient in No. 3: this is the geometric average or (G. A.) of the series.

This method of procedure permits us to state the process for computing the geometric mean in the same general form as that

for computing the arithmetic average. Using the data in the illustration above, the procedure is as follows:

$$\text{Log. G. A.} = \frac{\log. 8 + \log. 7 + \log. 5 + \log. 4 + \log. 2 + \log. 3}{6}$$

This gives only the logarithm of the geometric average, from which by the use of a table of logarithms we determine the geometric average.

The Weighted Geometric Average.—It may be necessary to assign weights to the several items in the group. In this case, the situation requires the computation of what may be called the weighted geometric average. In its fundamental elements, this process is not different from that followed in determining the simple geometric average. In making this computation instead of multiplying the several items in the group by their respective weights as we do in computing the weighted arithmetic average, we raise the several items to powers indicated by their respective weights and then continue as in the case of the simple geometric average. Assigning arbitrary weights to the items in the illustrative group above, the procedure is as shown in Table 1 below.

Numbers	Weights	Numbers Weighted	Logs of Weighted Numbers
8	3	512	2.70927
7	2	49	1.69020
5	4	625	2.79588
4	3	64	1.80618
2	6	64	1.80618
3	5	243	2.38561
Total	23		13.19332

TABLE 1

$$\text{Hence, Log. G. A.} = \frac{13.19332}{23} = .57362$$

$$\text{And, the G. A.} = 3.74645$$

The distinctive feature of this process is the presence of weights but so far as the process of computation is concerned these

weights mean nothing more than that each of the items in the group are to be taken a certain number of times. This will be clarified by expressing the illustrative series as 8, 8, 8, 7, 7, 5, 5, 5, 5, 4, 4, 4, 2, 2, 2, 2, 2, 2, 3, 3, 3, 3, 3. Thus, each item is taken as many times as is indicated by its weight and the procedure becomes identical with that employed in computing the simple geometric average.

The Use of Geometric Average.—Generally speaking, wherever the rate of change is involved the geometric average will be required. This is typified by a problem which has been met frequently in practical work since the war—viz., the average annual rate of increase in rates of pay from 1913 to some post-war period. To simplify the computation, suppose the average rate for typists increased 30% from January 1917 to January 1920, a period of 3 years. What was the average annual increase?

The apparent method of procedure would be to divide 30 by 3, but the fallacy of this method can be readily established simply by testing it out. It is the geometric rather than the arithmetic process that is involved and consequently the proper method of procedure is to compute the geometric average. That is, the correct result will be obtained by the following process:

$$\begin{aligned}\text{Av. annual rate} &= \sqrt[3]{.30+1} - 1\\ &= \sqrt[3]{1.30} - 1\\ &= 1.091 - 1\\ &= 0.91 \text{ or } 9.1\%\end{aligned}$$

Thus the average annual percentage increase is 8.2% as against 13 1/3 obtained by dividing 120 by 9.

Likewise in estimating the population for intercensus periods the geometric average is used, the assumption being that the rate of increase is the same for each of the intercensal years. Or if we wish to know the average annual percentage increase for a period of years in our capital and plant investment, building investment, and so forth, the geometric average will be required.

Many similar problems will arise in practical work in which the geometric average is the only average that can be used, such

as computing the average rate of increase in fixed capital, gross revenue, depreciation reserves, number of employes, and so forth. Its use, however, is confined almost exclusively to problems of this type.

The Harmonic Average—Computation.—The harmonic mean may be descriptively defined as the reciprocal of the arithmetic average of the reciprocals of the several items in the group. The method of computing it is simple in meaning but, as in the case of the geometric average, it is somewhat difficult to accomplish in practical work. Following is the specific method of procedure:

1. List the reciprocals of the items in the group—*i.e.*, express each item as a denominator of unity—½ and ¼, for example, being the reciprocal of 2 and 4 respectively.
2. Take the arithmetic average of these reciprocals—*i.e.*, summate the fractions and divide by their number.
3. Take the reciprocals of the result obtained in No. 2; this is the harmonic average.

The following simple illustration will clarify the process:

Let us take the group of data, 2, 4, 6, 8, 10 and 12.

The first step noted above results in:

$$\tfrac{1}{2} + \tfrac{1}{4} + 1/6 + 1/8 + 1/10 + 1/12$$

The second step in,

$$(\tfrac{1}{2} + \tfrac{1}{4} + 1/6 + 1/8 + 1/10 + 1/12) \div 6$$

$$\text{or } \left(\frac{60}{120} + \frac{30}{120} + \frac{20}{120} + \frac{15}{120} + \frac{12}{120} + \frac{10}{120}\right) \div 6$$

$$\text{or } \frac{147}{120} \div 6$$

$$\text{or } \frac{147}{720}$$

The third step in,

$$1 \div \frac{147}{720}$$

$$\text{or } \frac{720}{147}$$

or 4.9, the harmonic average

The weighted harmonic average may be computed simply by applying the respective weights, to the reciprocals before sum-

mating them. This is equivalent to repeating the fraction as many times as there are units in its weight. As in the case of the weighted geometric and arithmetic average the weighted process fundamentally is identical with the simple process.

The short cut method outlined in Appendix E may be applied to the computation of the harmonic average. It is clear that the second step is nothing but the computation of the arithmetic average and consequently the process may be treated in every way just as the process for determining the arithmetic average.

The Use of the Harmonic Average.—While the use of the harmonic mean is much more limited in practical work, there are cases where it is required. Its chief use in business management is in determining average performance. It may also be required in problems involving prices stated in terms of so much of a commodity for a *unit of money* or so much money for a *unit of the commodity.*

References

Bowley, Arthur L. Elements of Statistics, 2nd Ed., 1902, Ch. V.
—— An Elementary Manual of Statistics, 1915, Ch. III.
Davies, George R. Economic Statistics, 1922, pp. 20-29.
Forsyth, C. H. Mathematical Analysis of Statistics, 1924, 79-93.
Kelley, Truman L. Statistical Method, 1922, Ch. III.
King, Willford, I. Elements of Statistical Method, 1912, Ch. XII.
Jerome, Harry. Statistical Method, 1924, Ch. VII.
Mills, Frederick C. Statistical Methods, 1924, Ch. IV.
Riegel, Robert. Elements of Business Statistics, Chs. X and XI.
Rietz, H. L. Handbook of Mathematical Statistics, 1924, pp. 23-27, 32.
Rugg, H. O. Statistical Methods Applied to Education, 1917, Ch. V.
Secrist, Horace. An Introduction to Statistical Methods, 1917, Ch. VIII.
West, C. J. Introduction to Mathematical Statistics, 1917, Ch. IV.
Yule, G. U. An Introduction to the Theory of Statistics, 1919, Ch. VII.
Zizek, Franz. Statistical Averages (Translation by Persons), 1913, Parts I and II.

CHAPTER XL

RELIABILITY OF THE AVERAGE—DISPERSION AND SKEWNESS

Measuring Variations from the Average.—The average is a single figure taken as typical of the group. Since the same average value may result from series of data which differ widely in detail, we want to know in any given case how well it represents the data for which it stands. This comprehends a survey of the different methods of measuring the variation of the detailed facts of the group from their average and the discussion of how to compute these measurements. Two considerations are involved—(1) the measurement of the variation of the several items from the average, known as dispersion, and (2) the measurement of the comparative distribution of the items above and below the average, known as skewness. These measurements may be taken in absolute or relative form.

Absolute measurement is the determination of the numerical variation regardless of the size of the items: it is evident that the size of the average is important in this form of measurement. Relative measurement shows the variation related to the size of the items: it is the result obtained by expressing the variation in the form of percentage. Obviously, the relative form is generally required in making comparisons between two groups of data and even in the single group it is frequently preferable.

DISPERSION

General Conception.—The term "dispersion" is used to express the variation of the several items in a group from their average, and is designed to reflect the representativeness of the average. When the dispersion is slight the several items in the group are not widely different and the average well represents the group. When it is large, the several items differ widely and the average does not well represent them.

It is important to note that in dispersion we do not consider whether the separate items of the group are greater or less than the average: we merely determine the difference between the average and the several items without regard for the relative size of the two. For example, in Table 1 below the difference between the average and 68 is the same as the difference between the average and 70, and although in the former case the average is the greater and in the latter it is the less, yet the differences have the same significance in the computation of the dispersion of the series. In other words, in dispersion we undertake to measure the variations from the average regardless of algebraic signs. Dispersion takes account of the range and detailed distribution of the data in the group without any regard for the symmetry of the distribution.

While the foregoing is the conception of dispersion generally held, Professor Zizek (page 266 and following of his Statistical Averages) advances the idea that it is more serviceable to compute the dispersion of the items below the average and of those above the average separately.

Methods of Measuring Dispersion.—Many measures of dispersion have been suggested, many of which are of little interest to us in elementary business statistics. Hence, we shall consider only four methods (1) the range, (2) the quartile deviation, (3) the average deviation, and (4) the standard deviation. Others such as the percentile, mean difference, geometric, and so forth, are treated in the references listed on page 448.

The Range.—Since the *range* indicates the limits within which all of the items in the group lie, it may be taken as a measure of dispersion. This measurement involves merely a simple comparison of the two extreme items in the group—*i.e.*, the highest and the lowest. On this basis we have the following measures of absolute variability or dispersion:

1. The *range per se*—In Table 1 below, for example, the group of items range from a lower limit of 65 to an upper limit of 73. This gives us some information as to how the items in the group are "scattered" or dispersed. It may be supplemented by the figures noted in 2 and 3.

2. The difference between the extreme items. In Table 1, this is 73 minus 65, or 8.
3. The difference between the extreme items and the average. In Table 1, the difference between the lowest item and the arithmetic average is 4, and between the arithmetic average and the highest item 4. Thus, the values in the group extend below the average a distance of 4 and above the average, a distance of 4. Similar computations may be developed using the median or the mode.

HEIGHT OF A GROUP OF SALESMEN.
To the Nearest Inch.

Height in Inches	Number of Salesmen (Frequency) (f)	Deviations (d)	Deviations Squared (d^2)	Deviations Multiplied by Frequencies (fd)	Deviations Squared Multiplied by Frequencies (fd^2)
a	b	c	d	e	f
65	40	−4	16	160	640
66	35	−3	9	105	315
67	61	−2	4	122	244
68	81	−1	1	81	81
69	51	0	0	0	0
70	106	+1	1	106	106
71	51	+2	4	102	204
72	40	+3	9	120	360
73	40	+4	16	160	640
Total	505			956	2590

TABLE 1.—Illustrating the Method of Computing the Average and Standard Deviation.

The arithmetic average is 69+.

$$\text{The average deviation} - \text{Absolute} = \frac{956}{505} = 1.893$$

$$\text{Relative} = \frac{1.893}{69} = 0.0274$$

$$\text{The standard deviation} - \text{Absolute} = \sqrt{\frac{2590}{505}} = \sqrt{5.128} = 2.26$$

$$\text{Relative} = \frac{2.26}{62} = 0.033$$

It is frequently necessary to convert the absolute measurements of dispersion to the *relative* form, particularly when comparison is to be made between the dispersion of two or more separate groups. This is accomplished with the measurements noted above as follows:

1. Express the difference between the highest and lowest items as a percentage of (1) their sum, or (2) the average

of the series, preferably the former. From the data in Table 1 these computations are (1) .058, (2) .116.

2. Express the difference between the highest item and the average and the lowest item and the average separately as percentages of the average of the series. From the data of Table 1, these computations are respectively for the highest and lowest, 0.058 and 0.058.

These relative measurements are called coefficients of dispersion. It is generally preferable to use the average as the base. The relative measurements indicated above computed on the basis of the average, are very helpful when taken in connection with the range *per se*.

Generally speaking, the range is unreliable as a measure of dispersion. Its principal weakness is to be found in the fact that it reveals nothing as to the distribution of the details within the limits indicated by the extreme items—it is too greatly affected by extreme or unusual items. The items may vary in all sorts of ways within the range and not affect the range at all. Furthermore, by dropping the two extreme items the range may change materially. As a matter of fact since nearly all of the series in practical work are samples of greater or less completeness, the range will actually depend to a great extent on the number of items in the group: generally as we increase the number of items, we enlarge the range. On the whole, the range as an independent measure of dispersion is of little value in practical work. When properly associated with certain computations, however, the range is serviceable as is pointed out on pages 444 and 445.

The Quartile Deviation. (Q. D.)—The quartiles of the series are associated with the distribution in somewhat the same general way as the range; they fluctuate in value as the character of the distribution changes. Hence, the difference between the first and third quartiles may be taken as the basis for a second measure of dispersion. This measurement is sometimes called the "semi-interquartile range." We shall, however, adopt the term "quartile deviation."

The method of computing the quartile deviation is simple. The absolute variability or dispersion is determined by taking one-half of the difference between the upper and lower quartiles, and the

relative by dividing one-half the difference between the quartiles by their sum or by an average. In formula, letting Q_1 and Q_3 represent the lower and upper quartiles repectively:

$$\frac{Q_3 - Q_1}{2} = \text{Absolute quartile deviation}$$

$$\frac{\frac{Q_3 - Q_1}{2}}{\frac{Q_3 + Q_1}{2}} \text{ or } \frac{Q_3 - Q_1}{Q_3 + Q_1} = \text{Relative quartile deviation or the coefficient of dispersion}$$

The method of determining the quartiles is outlined on page 413.

Following shows the computation of these measurements for the data represented in Table 1 above, of which roughly Q_1 is 67 and Q_3 is 71:

$$\frac{Q_3 - Q_1}{2} = \frac{71 - 67}{2} = 2 = \text{Absolute quartile deviation}$$

$$\frac{Q_3 - Q_1}{Q_3 + Q_1} = \frac{71 - 67}{71 + 67} = 0.029 = \text{Relative quartile deviation or coefficient of dispersion}$$

The computation by the use of the average is apparent from the above.

It should be noted that the real basis of this measurement of dispersion is the difference between the quartiles and the median. The distance between the two quartiles includes one-half of the items in the series and in the perfectly symmetrical distribution the lower quartile is as far below the median as the upper quartile is above it. In such case, it is obvious that the difference between the median and either quartile is a definite reflection of the distribution about the median of the items within the quartile range (the middle half of the items in the series)—*i.e.*, it is a measure of the dispersion of these items. In the asymmetrical distribution, however, the difference between the median and the lower quartile is not equal to the difference between the median and the upper quartile. Consequently, we take the average of the two in order to get the average dispersion from the median.

It is clear from what has been said that the efficacy of this measurement rests on the distribution of the middle half of the items in the series. It is compared with other measurements of dispersion on pages 444 and 445.

The Average Deviation (A. D.)—A third method of meas-

uring dispersion is based on the actual variation of each individual item from the average of the group. This is perhaps the simplest and most perspicuous method. It consists simply in taking the difference between the average and each individual item and then taking the average of these differences, disregarding signs.

We may use the arithmetic average, the median, or the mode of the group in computing the average deviation, but in practical work we invariably use the arithmetic average. Likewise, in averaging the deviations we may compute the arithmetic average, the median, or the mode: but here again, in practical work we use the arithmetic average.

It should be emphasized that in this process all deviations are considered positive. We are interested merely in measuring the distance of the several items from the average and not in their positions with respect to the average. Hence, the question as to whether any particular item is greater or less than the average does not enter into the computation. This is important in connection with the consideration of the skewness of the distribution as indicated on page 445 and following but not in determining dispersion.

Specifically, the computation of the average deviation consists of the following steps:

1. Compute the average of the series (mean, median, or mode) from which the deviations are to be taken.
2. Take the difference between this average of the series and the several individual items as shown in column "c" of Table 1 on page 438.
3. Compute the average of these differences (usually the arithmetic): this is the average deviation. In the grouped series the differences will, of course, be weighted by the frequencies.

Letting "d" represent the difference between the average and a single item, "n" the total number of items in the group, "f" the frequencies, and "A. D." the average deviation, the formula for the computation is:

1. In simple frequency series, A. D. $= \frac{\Sigma d}{n}$
2. In a grouped frequency series, A. D. $= \frac{\Sigma fd}{n}$

In the grouped frequency series we make the same assumption in determining the deviations that is made in computing the arithmetic average of such a series—*i.e.*, we take the difference between the average and the mid-points of the individual classes as shown in Table 1 above. The classes are designated by the mid-points in this series.

The method of converting this absolute measurement to a relative one is accomplished by expressing the average deviation as a percentage of the average of the group which was used as a standard—that is,

$$\text{Coefficient of Dispersion} = \frac{\text{A. D.}}{\text{Average}}$$

Applied to the data given in Table 1 above, the coefficient of dispersion is 0.0274. This means simply that the individual items vary 2.74% on the average from the average of the group.

The application of the A. D. is discussed on page 444.

The Standard Deviation (S. D.).—A fourth method of measuring dispersion—the standard deviation—is likewise based on the differences between the average of the series and the several individual items contained in the series. It differs from the A. D. primarily in method of treating these differences. Whereas in the average deviation the differences are averaged directly, in the standard deviation the individual differences are squared before they are totaled.

While other types of averages may be employed in computing the standard deviation, the use of the arithmetic average is so generally followed that we shall confine our treatment to the process of computing the standard deviation from this average. It should be noted in this connection that the sum of the squares of the deviations from the arithmetic average is a minimum.

The process of computing the standard deviation is as follows:

1. Compute the arithmetic average of the series.
2. Take the difference between this average and the several individual items as shown in column "c" of Table 1 above.
3. Square each of these individual differences as shown in column "d" of Table 1 above.
4. Summate these squares and divide by the number of individual

items or frequencies in the group. In the grouped series the squares of the differences will be multiplied by their respective frequencies before being totaled.

5. Extract the square root of the quotient obtained in No. 4. The result is the standard deviation.

Again, letting "d" represent the differences, "f" the frequencies in the several groups, "n" the total number of items or frequencies in the series, and "S. D." the standard deviation, the formula for this computation may be expressed as follows:

1. In the simple frequency series,

$$\text{S. D.} = \sqrt{\frac{\Sigma d^2}{n}}$$

2. And, in the grouped frequency series,

$$\text{S. D.} = \sqrt{\frac{\Sigma f\, d^2}{n}}$$

Table 1 illustrates the method of computing the standard deviation in the grouped frequency series, where the S. D. is 2.26.

The measure of relative dispersion or the coefficient of dispersion is determined in the same manner as in the case of the average deviation—*i.e.*, by expressing the standard deviation as a percentage of the arithmetic average of the group. Thus,

$$\text{Coefficient of Dispersion} = \frac{\text{S. D.}}{\text{Arithmetic Average}}$$

Following this method of computation, the coefficient of dispersion on the basis of the standard deviation is 2.26 divided by 69, or 0.033 for the data in Table 1. The use of the S. D. is discussed on the following page.

The Use of Graphics.—A general picture of dispersion may be portrayed by means of graphic devices, particularly the frequency graph and the Lorenz curve. In the frequency graph the "spread" of the curve reveals dispersion; if the items are pretty well scattered throughout the series, the dispersion will be greater as shown in Figure 101, page 379, and if they are concentrated at a single point, the dispersion will be less as shown in "B" of Figure 106, page 396.

The Lorenz curve is used but little in practical work. Its

nature and the method of constructing it are discussed in the references listed at the close of this chapter.

How to Use Measures of Dispersion.—Thus we have noted four different measures of dispersion—the range, the quartile deviation, the average deviation, and the standard deviation. The matter of final importance is the practical application of these measurements. What concrete service do they render in practical business statistics?

In most instances in practical work the lone average without supplemental computations will suffice. Notwithstanding the rather extensive discussion of the subject given above measures of dispersion are used comparatively little. It is advisable, however, to check the representativeness of the average by inspection, and frequently by computing measures of dispersion. Thus, it may be well in some cases when presenting the average rate of pay, for example, or the average price, to state that the average variation of the individual rates in the one case or prices in the other from the average of the group is, say $2 per week or 15 cents a bushel or, relatively, 10% or 25% of the average of the group. The same practice should be followed here as was used in determining the scale in graphics—*i.e.*, the representativeness of the average should be determined by comparing the absolute dispersion with the smallest fluctuation of importance in the data or by comparing the coefficient of dispersion with the percent tolerance, so to speak, that can be allowed in the use of the data.

In making comparisons between series the relative dispersion or coefficient will be required. This is merely expressing the variations in percentage form, in order to make them comparable. Relative dispersion is required not only when units are different, as rates of pay with periods of service, but also when the units are the same but the averages differ substantially in magnitude.

The Particular Measure to Use.—On the whole the A. D. is preferable for general use in elementary business statistics just as the arithmetic mean is the preferable average. For some purposes, especially where rough, quick measurements are desired, the *range* and *quartile deviation* may be preferable. Thus, in the analysis of rates of pay we may say that they range from $15 to

$25 a week and are grouped generally toward the upper or lower limit or at the mid-point. Or, we may say that the middle-half of the rates lies between $18 and $20 and that they range from $15 to $25. In this way both the range and the Q. D. have a practical value.

In other cases, the S. D. may be preferable particularly in the more elaborate studies. As a matter of fact, it may be advisable to compute the Q. D., the S. D., and the A. D. The matter of prime importance is that we give thorough consideration to the facts of the specific case and follow the procedure which will serve the specific purpose. But unless special considerations require another type, the A. D. should be used. It is simple in computation, easily understood, and generally serves well the purposes of business statistics.

SKEWNESS

Symmetrical Distribution of Items Rare.—Dispersion has to do with the deviations of the items from a central tendency. It does not tell us how the items group themselves above and below this central tendency. It is obvious, therefore, that we must supplement the measure of dispersion with another which will reflect the relative distribution of the items above and below the average. In other words we must measure the *asymmetry* or *skewness* of the distribution.

A symmetrical series is one in which the distribution on either side of the average is similar in detail. It will be recalled that in this type of series, the mode, the median and the mean are equal. Now, skewness indicates the degree to which the distribution is pulled from the symmetrical form: it has to do with symmetry alone. In the perfectly symmetrical form the skewness is zero and as the distribution varies from the symmetrical, the skewness increases. Rarely in practical work do we find a symmetrical distribution and consequently we compute the skewness in order better to evaluate the series. The problem is to determine how much more the items are grouped on one side of the central tendency than on the other. Figure 101, page 379, represents a skewed distribution. If the skewness of a distribution is toward the lower limit of the series, it is negative: if toward the upper limit, it is positive.

There are two principal methods of measuring the skewness of a series of statistical data. The one is based on the sequence of the three averages—the mode, the median, and the mean. The other is based on the relation between the median and the first and third or upper and lower quartiles.

Variation in the Relative Values of Certain Averages.—In a symmetrical distribution the mode, the median, and the mean are equal. In the asymmetrical or skewed distribution they are unequal. The mode remains constant at the point of concentration, the median is pulled away from this point in the direction of the skewed portion of the series, and the mean is pulled still farther in the same direction. More specifically, in the moderately skewed distribution the median is pulled approximately two-thirds as far as the mean: in other words the difference between the median and the mode is equal approximately to two-thirds of the difference between the mean and the mode. Obviously, therefore, the difference between these averages may be taken as a measure of skewness. On this basis, absolute skewness is indicated simply by the difference between the mode and the mean. That is,

$$\text{Absolute skewness} = \text{mean} - \text{mode}$$

If the result of this computation is positive, the skewness is positive: if negative, skewness is also negative. In the light of the relative value of the median and mean in the moderately skewed series, approximately the same result may be obtained by taking three times the difference between the median and the mean. That is,

$$\text{Absolute skewness} = 3\ (\text{mean} - \text{median})$$

It will be noted that this process avoids the necessity of computing the modal value.

This absolute measurement is converted to the relative form by expressing it as a percentage of either the average deviation or the standard deviation. The latter is more generally used. Hence,

$$\text{Coefficient of skewness} = \frac{\text{mean} - \text{mode}}{\text{S. D.}}$$

or

$$\frac{3\ (\text{mean} - \text{median})}{\text{S. D.}}$$

Relation Between Median and Quartiles.—In the symmetrical distribution the median lies just half-way between the lower and upper quartiles. Skewness alters this relationship. The difference between the median and the quartile on the extended side of the distribution is greater than that between the median and the other quartile. Hence, we may take this as a rough measure of the skewness of the series.

Absolute skewness is measured simply by comparing the difference between the median and the lower quartile with the difference between the median and the upper quartile. In formula,

$$\begin{aligned}\text{Absolute skewness} &= (Q_3 - \text{median}) - (\text{median} - Q_1)\\ &= Q_3 + Q_1 - 2\ \text{median}\end{aligned}$$

Here, also if the result is positive, the skewness is positive—*i.e.*, the series is skewed toward the upper limit; and *vice versa.*

Relative skewness on this basis is determined by dividing the absolute measurement by the difference between the quartiles. Thus,

$$\text{Coefficient of skewness} = \frac{Q_3 + Q_1 - 2\ \text{median}}{Q_3 - Q_1}$$

Like the Q. D. this measurement of skewness does not take into account the extreme items of the group and for this reason it is not generally satisfactory.

Use of Measures of Skewness.—Measures of skewness are not used extensively in practical work. We can, however, derive some general service from such measurements. It is frequently of value to know that the skewness of a series is positive or that it is negative. In such cases we say that the average is so and so and that such and such per cent of the items are above it and below it respectively.

The statistician must use his judgment as to whether he shall compute the measurements of skewness, as of dispersion, in any given case. In most cases where the average is used they will not be required and wherever they are computed he should make sure that they offer real, concrete, services. Further discussion of them may be found in the references below.

References

Davies, George R. Economic Statistics, 1922, 29-44.
Forsyth, C. H. Mathematical Analysis of Statistics, 1924, pp. 94-102.
Jerome, Harry. Statistical Method, 1924, Ch. IX.
Kelley, Truman, L. Statistical Method, 1923, Ch. IV.
King, Willford I. Elements of Statistical Method, 1912, Chs. XIII, XIV.
Mills, Frederick C. Statistical Methods, 1924, Ch. V.
Riegel, Robert. Elements of Business Statistics, Chs. XII and XIII.
Rietz, H. L. Handbook of Mathematical Statistics, 1924, pp. 27-32, 33.
Rugg, H. O. Statistical Methods Applied to Education, 1917, Ch. VI.
Secrist, Horace. An Introduction to Statistical Methods, 1917, Ch. XI.
Yule, G. U. An Introduction to the Theory of Statistics, 1919, Ch. VIII.
Zizek Franz. Statistical Averages (Translation by Persons), 1913, Part III.

CHAPTER XLI

INDEX NUMBERS

Extensive Use of Index Numbers.—The extensive use of index numbers in business management today has given great impetus to the consideration of the practical aspects of the subject. We are realizing more and more the service to be derived from this method of analysis in the management of the business organization and are continually extending its application. Whereas in the past the use of this device was confined for the most part to the analysis of prices, it is being extended now to all sorts of data—personnel, financial, and operations. Thus we have indexes of wage rates, employment, sales, production, imports and exports, building materials, rents, and other important groups of items.

Coincident with the application of the method to other types of data more attention has been given to the development of series to serve specific purposes, such as those reflecting the trend of prices of a particular class of commodities or the wages of a particular class of workers, as contrasted with series designed to reflect the course of the general price or wage level. Series from internal data of a business concern are generally developed for the specific purpose of aiding in the solution of practical problems of the company.

The Nature and Purpose of an Index Series.—The meaning of an index number as implied in the dictionary definition of the term—*i. e.*, to indicate or point out—is too often disregarded, particularly in popular interpretations. Thus, the wage earner will apply the cost of living index, which is the composite result of a great variety of factors, to a specific item in his family budget and on the basis of this comparison, declare the index series erroneous. If the index number indicates that the cost of living has decreased over a period of time, the wage earner will make these improper comparisons and rejoin, knowingly, "Find it."

Now, by its very nature, the index number neither indicates nor

implies such a fact. Over any given period of time fluctuations of the several related items which make up the series, have varied; some have fluctuated in one direction and others in another and the index is designed to indicate or reflect the general trend of the composite. This is the essential nature of the index number and it must be held to strictly in the use of this method of analysis. Even if we were developing an index number of a single variable item—say, the price of a 12-ounce loaf of bread or the weekly wage of machinists—the course of the index series would not necessarily correspond minutely to the movement, over the same period, of the price paid by a single individual or the wage rate of a single machinist. The index number is not designed to reflect the course of a single fact over a period of time. It indicates only a general trend which is the composite of many fluctuations differing in direction and degree. Any single item may well fall on either side of this general trend. Further consideration is given to this point on page 461 in the discussion of the interpretation of the index number.

Fundamentally, the method of index numbers is a device designed to reduce different series of data to comparable bases. Its use greatly facilitates the following:

1. Comparison of the individual items in a single series.
2. Comparison of the trends of two or more series of items.
3. Combination of different series of related items to reflect the general trend of the group.

Methods of Computation.—In building up an index series what we really seek to do is to combine all the factors of importance in such way that each will influence the result in proportion to its relative importance. Because of the variation in the units in which the items are expressed and their relative importance, we cannot properly take the simple sums of actual data. Obviously, the attempt to combine butter at 40 cents a pound with steel billets at $40 a ton, or wages of a laborer at 45 cents an hour with the salaries of tool-makers at $75 a week, would lead to serious error. Likewise, it is probable that in any index of prices or wages in practical work the importance of these items would

vary greatly. The same is generally true for all kinds of data and hence we must resort to other methods.

A series of index numbers may be developed in a number of different ways chief among which are (1) average of relatives, (2) aggregates of actual items, and (3) relatives of aggregates.

Professor Irving Fisher has selected an "ideal" formula for the construction of an index series as a result of having subjected the various methods to two essential tests—first, the *time-reversal test* which requires that a shift in the base period should not change the relative values of the index figures for the respective periods, and second, the *factor-reversal test* which requires, in the price index, for example, that separate index series of the prices of the commodities and the quantities marketed when multiplied item by item should give the index for the value which is the product of the two. The "ideal formula" which provides for the combination of two separate series by means of the geometric average is as follows (Formula 353, pages 220 *ff* and 482):

$$P_1 = \sqrt{\frac{\Sigma\, p_1 q_0}{\Sigma\, p_0 q_0} \times \frac{\Sigma\, p_1 q_1}{\Sigma\, p_0 q_1}}$$

In which P_1 = the index number at a given period
p_1 = the price of any commodity at a given period
q_1 = the quantity of the same commodity at a given period
p_0 = the price of the same commodity at the base period
q_0 = the quantity of the same commodity at the base period

Professor Fisher suggests a modification of this formula as more practicable (Formula 2153, pages 428 and 488):

$$P_1 = \frac{\Sigma\,(q_1 + q_0)\, p_1}{\Sigma\,(q_1 + q_0)\, p_0}$$

This "ideal" formula is developed fully in Professor Fisher's "The Making of Index Numbers." This is a most valuable treatise on the subject of index numbers and should be thoroughly examined by anyone who contemplates the construction of complex index series.

Link relatives may be used when the primary comparison is successive years. In this method, the preceding year is the base for computing the index for any given year. A chain index is computed from these link relatives by multiplying them together serially. For the purposes of this book, this method has a com-

Average Wholesale Prices

Years	FOOD						CLOTHING						FUEL AND LIGHTING					
	Bread—Loaf Price per Lb. (New York)		Meat—Fresh Beef—Good Native Steers (Chicago)		Milk—Fresh Per Quart (New York)		Boots & Shoes Men's Vici Kid per Pair		Suitings—Serge—per Yard (Fulton Mills)		Ginghams—Amoskeag per Yard		Bituminous Coal—Kanawha Run of Mine f.o.b. Cincinnati per Short Ton		Anthracite—Egg per Long Ton		Petroleum—Crude Kansas-Oklaho. per Barrel	
	Absolute	Relative	Absolute	Relative	Absolute	Relative	Absolute	Relative	Absolute	Relative	Absolute	Relative	Absolute	Relative	Absolute	Relative	Absolute	Relative
a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s
1913	$0.042	100.0	$0.130	100.0	$0.044	100.0	$2.867	100.0	$1.131	100.0	$0.065	100.0	$2.200	100.0	$5.064	100.0	$0.934	100.0
14	.044	104.0	.136	105.3	.042	93.9	2.983	104.1	1.078	95.4	.063	96.2	2.200	100.0	5.059	99.9	.798	85.5
15	.048	112.0	.129	99.5	.042	94.8	3.104	108.3	1.194	105.6	.064	98.2	2.200	100.0	5.046	99.7	.583	62.4
16	.048	112.5	.138	106.7	.045	102.3	3.504	122.2	1.513	133.8	.081	125.1	2.675	121.6	5.291	104.5	1.258	134.7
17	.069	163.4	.167	129.1	.063	143.1	5.067	176.7	2.273	201.0	.128	196.8	4.583	208.3	5.622	111.0	1.775	190.0
18	.073	171.9	.221	170.9	.078	174.9	5.442	189.8	3.391	300.0	.180	277.2	3.883	176.5	6.509	128.5	2.197	235.2
19	.074	175.5	.233	180.2	.084	190.5	7.246	252.8	2.989	264.4	.189	290.9	4.108	186.7	7.920	156.4	2.279	243.9
20	.087	205.7	.230	177.9	.085	191.4	8.338	290.8	3.679	325.4	.221	340.6	5.850	265.9	9.167	181.0	3.404	364.3

(Data from Bulletin 335 of Bureau of Labor Statistics.)

Table I.—Sample Data to Illustrate the Methods of Computing Index Numbers

The relative prices are determined by taking the average for 1913 as a base. Thus, each of the entries in column b are divided by .042 to obtain the entries in column c; those in column d are divided by 130 to obtain the entries in column e; and so on for the rest of the commodities. The relatives have been copied from the bulletin. The computation of the index series is completed in subsequent tables.

paratively limited use. It is fully discussed in the references given on page 463.

Average of Relatives.—This method of developing index numbers consists simply of computing series of relative numbers for separate groups of items and then combining these series of relatives into a single index series by averaging them item by item. Table 1 above together with the accompanying text will clearly illustrate the process. The detailed steps of procedure are as follows:

1. Compute relatives by expressing the individual numbers in each group of facts separately as percentages of the base (columns "c," "e," "g," "i," "k," "m," "o," "q," and "s" in Table 1 above).
2. Take the simple or weighted average, item by item, of these relatives to get the combined figures for the separate groups of items (Tables 2 and 3 below).
3. Again, take the simple or weighted average, item by item, of the group figures resulting from No. 2 (Table 2 below).

Aggregates of Actual Data.—This method involves simply the summation of the products obtained by multiplying the actual data by the weights assigned to them if weights are used, or, the summation of the data alone (Table 4 below) if weights are not assigned. Thus in price indexes we multiply the prices of each commodity represented by the quantities assigned to them as weights and summate the products as shown in Table 5 below; in an index of wage rates, we multiply the wage rates for the separate wage groups by the relative weights assigned to them based perhaps on the number of employes at each rate and total the products; and likewise for other classes of data. In each of these cases, it is assumed that the representative price or wage for each item at each date as well as the weights, are at hand. The selection of these figures and weights is a very difficult part of the work, however.

As may be determined from the references below, this method is used by Bradstreet and Dun although the detailed form is different in the two cases.

Average Wholesale Prices
Index Numbers by Average of Relatives

Years	FOOD GROUP				CLOTHING GROUP				FUEL AND LIGHTING GROUP				Average of Relatives for the Three Groups
	Bread—Relative	Meat—Relative	Milk—Relative	Average of Relatives	Shoes—Relative	Suitings—Relative	Ginghams Relative	Average of Relatives	Bituminous Coal Relative	Anthracite Coal Relative	Petroleum Relative	Average of Relatives	
a	b	c	d	e	f	g	h	i	j	k	l	m	n
1913	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
1914	104.0	105.3	93.9	101.1	104.1	95.4	92.2	97.2	100.0	99.9	85.5	95.1	97.8
1915	112.0	99.5	94.8	102.1	108.3	105.6	98.2	104.0	100.0	99.7	62.4	87.4	97.8
1916	112.5	106.7	102.3	107.2	122.2	133.8	125.1	127.0	121.6	104.5	134.7	120.3	118.2
1917	163.4	129.1	145.1	145.2	176.7	201.0	196.8	191.5	208.3	111.0	190.0	169.8	168.8
1918	171.9	170.9	174.9	172.6	189.8	300.0	277.2	255.7	176.5	128.5	235.2	180.1	202.8
1919	175.5	180.2	190.5	182.1	252.8	264.4	290.9	269.4	186.7	156.4	243.9	195.7	215.7
1920	205.7	177.9	191.4	191.7	290.8	325.4	340.6	318.9	265.9	181.0	364.3	270.4	260.3

(Data from Bulletin 335 of Bureau of Labor Statistics.)

Table 2.—Index Number of Simple Average of Relatives

The entries in column e are the index for the "food group" and are determined by taking the simple arithmetic average, year by year, of the entries in columns b, c, and d. Columns i and m, the index numbers respectively for the "clothing group" and the "fuel and lighting group" are determined in similar manner. Column h, which is the index series for the three groups combined, is determined by averaging, year by year, the entries in columns e, i, and m. The relatives in this table are taken from Table 1 above.

AVERAGE WHOLESALE PRICES.

Years	FUEL AND LIGHTING GROUP							
	Bituminous Coal—Per Ton (Weight=40)		Anthracite Coal—Per Ton (Weight=15)		Petroleum—Per Barrel (Weight=216)		Sum of the Products (c+e+g) (Sum of the Weights=271)	Index Group
	Relative	Product	Relative	Product	Relative	Product		
a	b	c	d	e	f	g	h	i.
1913	100.0	4,000.0	100.0	1,500.0	100.0	21,600.0	27,100.0	100.0
1914	100.0	4,000.0	99.9	1,498.5	85.5	18,468.0	23,966.5	88.4
1915	100.0	4,000.0	99.7	1,495.5	62.4	13,478.4	18,973.9	70.0
1916	121.6	4,864.0	104.5	1,567.5	134.7	29,095.2	35,526.7	131.1
1917	208.3	8,332.0	111.0	1,665.0	190.0	41,040.0	51,037.0	188.3
1918	176.5	7,060.0	128.5	1,927.5	235.2	50,803.2	59,790.7	220.6
1919	186.7	7,468.0	156.4	2,346.0	243.9	52,682.4	62,496.4	230.6
1920	265.9	10,636.0	181.0	2,715.0	364.3	78,688.8	92,039.8	339.6

(Data from Bulletin 335 of Bureau of Labor Statistics)

TABLE 3.—Index Numbers by Weighted Average of Relatives

The entries in columns c, e, and g are obtained by multiplying the corresponding series of relatives by the weights shown at the heads of the columns. The entries in column h, when divided, item by item, by the sum of the weights, results in the final index series listed in column i.

The weights used are taken from Bulletin 335 of the Bureau of Labor Statistics and represent as nearly as could be determined the total number of units (in millions) of the commodity sold in the markets in the year 1919.

AVERAGE WHOLESALE PRICES.

Year	Bituminous Coal (Short Ton)	Anthracite Coal (Long Ton)	Petroleum (Barrel)	Simple Aggregative Index (b + c + d)
a	b	c	d	e
1913	$2.200	$5.064	$0.934	$ 8.198
1914	2.200	5.059	0.798	8.057
1915	2.200	5.046	0.583	7.829
1916	2.675	5.291	1.258	9.224
1917	4.583	5.622	1.775	11.980
1918	3.883	6.509	2.197	12.589
1919	4.108	7.920	2.279	14.307
1920	5.850	9.167	3.404	18.421

(Data from Bulletin 335 of Bureau of Labor Statistics.)

TABLE 4.—Simple Aggregative Index

The index numbers in column e are determined simply by taking the total year by year of the entries in columns b, c, and d.

AVERAGE WHOLESALE PRICES
Index Number by Method of Aggregates

Years	FUEL AND LIGHTING						Sum of the Products
	BITUMINOUS COAL—Per Net Ton (Weight = 40)		ANTHRACITE COAL—Egg—Per Gross Ton (Weight = 15)		PETROLEUM—Crude Per Barrel (Weight = 216)		
	Price per Ton	Product	Price per Ton	Product	Price per Barrel	Product	
a	b	c	d	e	f	g	h
1913	$2.200	88.00	$5.064	75.960	$0.934	201.744	365.704
1914	2.200	88.00	5.059	75.885	0.798	172.368	336.253
1915	2.200	88.00	5.046	75.690	0.583	125.928	289.618
1916	2.675	107.00	5.291	79.365	1.258	271.728	458.093
1917	4.583	183.32	5.622	84.330	1.775	383.400	651.050
1918	3.883	155.32	6.509	97.635	2.197	474.552	727.507
1919	4.108	164.32	7.920	118.800	2.279	492.264	775.384
1920	5.850	234.00	9.167	137.505	3.404	735.264	1,106.769

(Data from Bulletin 335 of Bureau of Labor Statistics)

TABLE 5.—The Weighted Aggregative Method of Computing Index Numbers

The actual prices in columns b, d, and f, are multiplied respectively by the weights shown at the column headings and the products totaled to give the final index series in column h. To get a combined index for the three groups of commodities shown in Table 1, a series similar to that of column h in this table would be computed for the other two groups and then these three series would be totaled year by year to get the combined series.

AVERAGE WHOLESALE PRICES

Year	Aggregative Index	Relatives with Aggregate for 1913 as Base
a	b	c
1913	365.704	100.00
1914	336.253	91.95
1915	289.618	79.19
1916	558.093	152.61
1917	651.130	178.05
1918	727.507	198.94
1919	775.384	212.03
1920	1106.769	302.64

TABLE 6.—The Method of Relatives of Aggregates

The entries in column b are converted to the relatives of column c simply by dividing each item by the aggregate for 1913 which has been taken as the base. Obviously the relatives with the aggregate for any other year as the base can be readily computed. (Data from column "h" of Table 5.)

Relatives of Aggregates.—It is clear that these aggregates can be converted to the relative form simply by dividing each one by the aggregate chosen as the base, as shown in Table 6 above. A form of this method is used by the Federal Bureau of Labor Statistics in their development of index numbers of wholesale prices.

Selection of the Items.—One of the most important as well as one of the most difficult steps in the construction of an index series is the selection of the items that are to be included—*i.e.*, the commodities in a price index, the occupational or wage groups in an index of wages, the products in a production series, or the units in other classes of data.

Broadly, we want to include in the index series data representing all the items which are important from the standpoint of the purpose of the index and we want these weighted according to their relative importance. The number and kinds of items to be included as well as the weights to be assigned will depend largely on the character and purpose of the index series. There are no hard and fast rules to guide us: it is largely a matter of judgment. The kinds of items to be included, their number, and the weights assigned to them can be determined only on the basis of the facts of the particular series being developed.

Bulletin 284 of the U. S. Bureau of Labor Statistics (pages 33 to 59) contains an excellent discussion of this point as related to price index series. An inspection of the different series listed in Appendix C will show the variation in the number of items used in currently published series.

Collecting the Data.—The selection of the data to be used in the construction of the index series is generally very difficult. Much depends on the way it is done. What has been said relative to the collection of statistical data will apply specifically in this case. The use of the data in the construction of an index number becomes the purpose of the collection or assembling of the data and the many problems which arise must be solved in the light of this purpose.

For the most part, the relatives or index numbers developed in practical work are simple in character and the data are gen-

erally obtained directly from the basic records of the statistical organization. These cases are generally confined to such problems as converting rates of pay, sales, output, transportation costs, selling prices, purchase prices, and so forth, to a relative basis in which the combinations are comparatively simple.

In the development of the more elaborate *general purpose* series a special collection is usually necessary and because of the volume of data generally required in this type of series as well as the many intricate problems involved in obtaining them, this becomes one of the most difficult phases of the work. The quotation on page 130 from Bulletin 284 of the Federal Bureau of Labor Statistics, gives a vivid picture of the problems involved in collecting data for an index number of wholesale prices.

It should be noted that in some cases it is necessary to collect data in establishing weights also and this may be even more difficult than the problem of collecting the principal facts.

Selecting the Base.—As stated on page 367, relatives consist merely of expressing the data of the series as percentages of a single figure in order to make them comparable. This figure is called the base of the series and the relative or index for the corresponding period equals 100. Whenever relatives are involved in the construction of index numbers—such as the average of relatives or relatives of aggregates—it is necessary to select a base.

It should be noted that usually the facts as of a certain selected time period are taken as the base and that in averages of relatives the base figure for any given item is the figure representing that item as of the time period chosen as the base. Thus, if in combining the wages of typists, stenographers, and senior clerks to get an index number of wages of experienced office help, the average for January 1913 is taken as the base, then in the average of relatives the wages of typists, for example, will be expressed as percentages of the average for typists in January 1913, and similarly for the other classes. This aspect of the method is illustrated by Table 1, page 452.

The problem of selecting the base is much more important in the method of *averages of relatives* than it is in *relatives of*

aggregates since in the latter case the base can be readily shifted, while in the former the shifting of the base will involve a repetition of the entire process or the use of the less accurate "short-cut" method noted below. There are no definite directions that can be given for the selection of the base. It may be the data for a single day, week, month, year or other time period, in total, or in the form of an average. Generally, in the average of relatives method we should select as base that particular figure with which comparisons are most significant. In the cost of living and price index series in the United States, for example, 1913 is used as the base because we are particularly interested in making comparison with a pre-war period.

As noted above, the base may be readily shifted in a series constructed from relatives of aggregates—Table 6 above. In a series developed through the method of averages of relatives, the only accurate way of shifting the base is to develop the series anew on the new base. This is so expensive in time and money that it is impracticable. As a substitute we use a so-called short method which consists merely of selecting as the new base one of the numbers in the series and dividing this into the other items in the series as shown in Table 7 below. While this method does not produce as accurate results as repeating the process, it is for the most part sufficiently accurate for practical purposes.

The foregoing discussion has been confined to the fixed base. We may also develop the so-called *chain index numbers* by taking the preceding year as the base in each instance. A comparison of the *chain* and *fixed-base* series is given on pages 81 to 91 of Bulletin 284 of the U. S. Bureau of Labor Statistics.

Selecting the Average.—A great deal has been written on the question of the proper average to use in constructing an index series through the average of relatives. The arithmetic mean, the median, and the geometric mean, have been championed and each has been used at one time or another in the actual construction of price index series. For practical purposes, it may be said, the arithmetic mean is preferable in most cases. While it is not so convenient as the geometric average from the point of view of shifting the base—and this is a serious disadvantage—nevertheless,

it is generally more familiar and is more easily computed. It is safe to say that the disadvantages of the median are such as to make its use inadvisable in practical business statistics except in very special cases. Finer points of comparison will be found on pages 68 to 81 of Bulletin 284 of the Federal Bureau of Labor Statistics.

INDEX NUMBERS OF WHOLESALE PRICES.
By Years, 1913-1923.
(ALL COMMODITIES)

Year	Average for 1913 taken as Base	Average for 1917 taken as Base
a	b	c
1913	100	56
1914	98	55
1915	101	57
1916	127	72
1917	177	100
1918	194	110
1919	206	116
1920	226	128
1921	147	83
1922	149	84
1923	154	87

(Source: *Monthly Labor Review*, February 1924, p. 81.)

TABLE 7.—Illustrating the short method of shifting the base of an index series. Each entry in column b is divided by 177, the entry for the new base. The results appear in column c.

Selecting the Weights.—"When an effort is made to ascertain the relative importance of the various commodities included, and to apply some plan by which each commodity shall exercise an influence upon the final results proportionate to its relative importance, the index number is said to be weighted. When, on the contrary, no such effort is made, but every commodity is supposedly allowed just the same chance to influence the result as every other commodity, the index number is said to be unweighted or simple."

This quotation, taken from page 59 of Bulletin 284 of the U. S. Bureau of Labor Statistics, sets forth the distinction between

what is generally known as a *weighted* and a *simple* series. As a matter of fact, however, as the bulletin proceeds to note, we do not get an unweighted series merely by dispensing with the use of definitely assigned weights. In the absence of definitely assigned weights we get an "unconscious" or "chance" weighting and the "real problem of the maker of index numbers is whether he shall leave weighting to chance or seek to rationalize it."

The particular weights to use in any given case can be determined only in the light of the facts of the particular situation, the purpose playing a vitally important part in this decision. In price index series, for example, we may use consumption, distribution, or production figures, in a wage series the number of employes, and in a sales series the relative profits on the different classes of goods. These weights may be taken for the base period, the current period, or any other time period. Professor Fisher's "ideal formula" involves the use of weights representing both the base and current period. This point is treated at length on pages 59 and 68 of Bulletin 284 of the U. S. Bureau of Labor Statistics, and in "The Making of Index Numbers" by Professor Irving Fisher.

Conclusion—The Interpretation of an Index Series.—In the interpretation of the index number, we must keep the nature of this method of analysis clearly in mind: namely, that generally the final figures which make up the index are the result of a composite of many movements and consequently cannot be taken as indicating the course of any single fact in the field which it represents. On page 294 of his "Principles of Economics," Vol. 1, Professor F. W. Taussig states:

> . . . it must be borne in mind that no index number corresponds to a real thing. It is not like the mean of certain observations in natural science—such, for example, as those for measuring the distance between the earth and the sun—of which any one may err, but whose average will point to a single specific fact. An index number points to no single fact. It gives, to repeat, only an indication of a general trend of prices. People often speak and think loosely on this topic, as if an index number told the whole story once for all. There is no one change in prices. There is a medley of many changes, different in direction and degree. All that we can

hope to secure by averaging and summarizing is some concise statement of the general drift.

The force of this statement will apply to index numbers in any field of data. It may be simply illustrated by Figure 111 below, in which six series of data separately and in combination are represented by the straight line curve. It will be noted from this chart that the "housing" group, for example, increased materially from the middle of 1920 to the middle of 1921, while the "food" group and "clothing" group decreased very greatly over the same period.

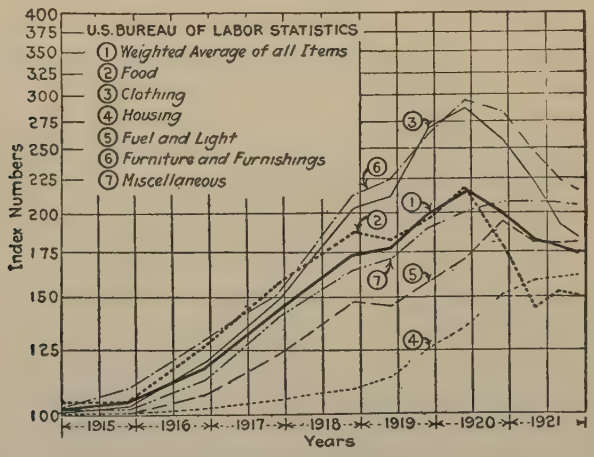


FIGURE 111.—The Cost of Living from 1915 to 1921 as Reported by the U. S. Bureau of Labor Statistics

The practical importance of the point may be brought out by examining the interpretation which is sometimes given, namely, that the index series is incorrect because it does not correspond to the movement of the price of a single commodity in a given locality. But the index series is not intended to reflect the movement of such detailed facts. There are two serious fallacies in this inference. In the first place, the effect of other factors may have been sufficient to throw the trend of the series in a direction opposite to the movement of these specific prices: the trend is a composite of different factors some moving down more than others

and perhaps some even moving upward, and the particular prices in question may be at variance with the composite. In the second place the specific prices in question may lag behind the general composite and consequently may not show a drop until some time after the general trend has dropped.

Furthermore, in the use of an index series we should thoroughly examine it from the standpoint of the contemplated purpose. We should know the items of which it is composed, the method of computation, and, if pertinent, the base period, the average used, and the method of weighting. In short, we should make sure that it is satisfactory for the purpose in mind.

In Appendix C, pertinent facts are listed relative to a number of index series of special interest in business management and similar facts for other series will be found in Bulletin No. 284 of the U. S. Bureau of Labor Statistics. This Bulletin contains a thorough and satisfactory treatment of the subject and should be in the hands of every teacher of statistics and available to students and also in the files of the business statistical organization.

References

Bowley, Arthur L. Elements of Statistics, 2nd Ed., 1902, Ch. IX.

Davies, George R. Economic Statistics, 1922, Ch. III (Wages and Prices), Ch. IV (Quantity).

Fisher, Irving. The Making of Index Numbers, 1922.

Karsten, Karl G. Charts and Graphs, 1923, Ch. XXVI.

Kelley, Truman L. Statistical Methods, 1923, Ch. XIII.

King, Willford I. Elements of Statistical Method, 1912, 178-185.

Jerome, Harry. Statistical Method, 1924, Ch. XI (Construction), Ch. XII (American Series).

Mills, Frederick C. Statistical Methods, 1924, Ch. VI (Prices), Ch. IX (Physical Volumes).

Rietz, H. L. Handbook of Mathematical Statistics, 1924, Ch. XII (Allyn A. Young).

Secrist, Horace. An Introduction to Statistical Methods, 1917, Ch. IX (Construction); Ch. XI (American Price Series).

U. S. Bureau of Labor Statistics. Bulletin No. 284—Index Numbers of Wholesale Prices in the United States and Foreign Countries.

CHAPTER XLII

STATISTICAL RELATIONSHIPS—CORRELATION

The Meaning of Correlation.—This discussion aims merely to point out what correlation is and how it may be serviceable in elementary statistical work in practical business management. For more complete treatises the reader is referred to the list of references on page 470.

In previous chapters we have discussed the analysis of a single series of data. Correlation considers the relationship between series. Between two series of variable data, it means that as one fluctuates there is a tendency for the other to fluctuate in the same or opposite direction. Thus stock prices and commodity prices, money rates and security prices, rates of pay and periods of service of employes, wholesale and retail commodity prices, and so forth, tend to move together and hence are said to be correlated. If the fluctuation in the one series is associated with a precisely proportional fluctuation in another, the correlation is said to be perfect, but rarely, if ever, do we have this condition in the field of practical business statistics. It will be observed from what has been said that correlation may be (1) *positive or direct* when the two series fluctuate in the same direction, or, (2) *negative or inverse* when they fluctuate in opposite directions. As wholesale prices increase, retail prices likewise tend to increase and the correlation is *positive*. On the other hand as prices decrease business failures tend to increase and we have *negative* correlation.

It should be noted that correlation does not necessarily imply causation. Two series may fluctuate together and yet the one not be the cause of the other. Causation does not exist without correlation but correlation may exist without causation. However, correlation suggests causation—it is a "tip" or "lead," so to speak, which assists us in eliminating factors which do not enter into the causal relationship and facilitates the location of the cause.

It should be noted also that correlation may exist between more than two series, in which case we have *multiple correlation*.

This is illustrated by the relationship between advertising, sales, and profits or between payroll, production costs, and profit. The individual relationships between the several pairs of series in a multiple correlation group, independent of the other variables, represent what is called *partial correlation.* The analysis of multiple and partial correlation is somewhat more complicated than the analysis of the relationship between two series and consequently

CHEST AND WAIST MEASUREMENTS.
(2,136 Women Students—Pratt Institute, Brooklyn, N. Y.)

EXPANDED CHEST IN Inches	WAIST IN INCHES								Number of Students
	20	22	24	26	28	30	32	34	
28	3	13	3	1					20
30	13	130	65	5	2				215
32	2	205	418	85	2				712
34		44	405	305	35	1			790
36			40	152	88	16	1	2	299
38			2	16	31	22	11	4	86
40					1	2	6	3	12
42							1	1	2
	18	392	933	564	159	41	19	10	2,136

TABLE 1.—Correlation Table

Data used by G. E. F. Saum, industrial engineer, in an article on "Standards in Clothing Manufacture," appearing on p. 59, *Management and Administration,* January, 1925.

It will be noted that the character of the table provides that the classes in the two series be paired and that the number in any given class are distributed by measurements in the other series. Thus, of the 20 students having a chest measurement of 28 inches, 3 have a waist measurement of 20 inches; 13 of 22 inches, 3 of 24 inches; and 1 of 26 inches.

will not be considered in these discussions. Those who are interested will find treatises on the subject in Chapter XII of Yule's "Introduction to the Theory of Statistics"; Chapter X of Forsyth's "Mathematical Analysis of Statistics"; Chapter XI of Kelley's "Statistical Method"; and in other books listed on page 270.

Measuring Correlation.—It is not enough merely to know that correlation exists between two series of facts; we want to get some indication of its extent. This may be done (1) by constructing the correlation table, (2) by associating the two series graphically, and (3) by the computation of a coefficient of correlation.

The Correlation Table.—By arranging the two series of data in double frequency table such as Tables 1 and 2 we get the *correlation table* from which we can detect the presence or absence of correlation and determine whether it is positive or negative. In these tables it is clear that as the items in the one series increase those in the other series tend also to increase. This indicates the presence of positive correlation but tells us only in a very rough manner the extent to which it exists. If the items in the correlation table are pretty well concentrated correlation is rather pronounced and, if not, correlation is less pronounced or even absent altogether.

CORRELATION TABLE.

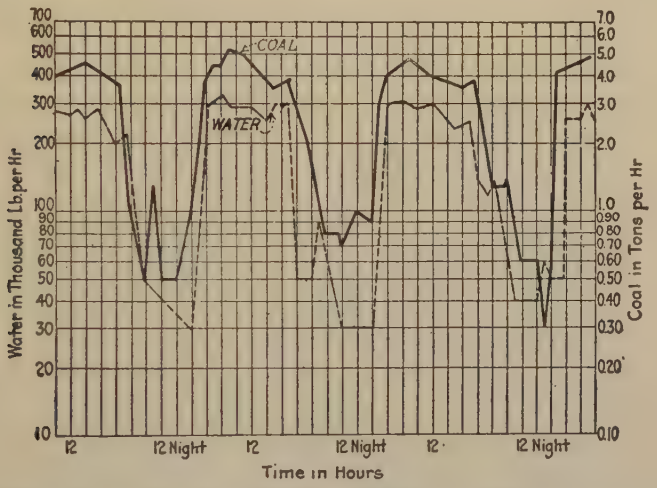
Rates of Pay and Period of Service of 790 Factory Workers.

Rates of Pay (Cents per hour)	Total Number in Group	TOTAL DISTRIBUTED BY PERIOD OF SERVICE								
		Under 1 Yr.	1 Yr. Under 2 Yrs.	2 Yrs. Under 3 Yrs.	3 Yrs. Under 4 Yrs.	4 Yrs. Under 5 Yrs.	5 Yrs. Under 6 Yrs.	6 Yrs. Under 7 Yrs.	7 Yrs. Under 8 Yrs.	8 Yrs. and Over
35–39	44	27	13	3	1					
40–44	63	34	19	10						
45–49	96	25	49	20	1	1				
50–54	102	10	52	22	10	6	2			
55–59	213	9	53	64	14	35	25	8	2	3
60–64	145	11	28	8	7	22	29	14	7	19
65–69	57	4	8	17	3	2	9	9	3	2
70–74	31		9	3	3	5	2	3	1	5
75 and over	39	.2	5	1		9	3	2	3	14
Total	790	122	236	148	39	80	70	36	16	43

TABLE 2.—Correlation Table in the Field of Personnel Statistics

Use of Graphs in Determining Correlation.—A more definite method of determining the correlation between two series of data is to plot them in simple graphic form by means of the straight line curve and then to compare the graphs. Correlation is designed to reflect the "similarity of behavior" or the "moving together" of the two series and since the curve graph clearly portrays the movement of the series, we can by inspection determine roughly whether the two curves *move together*—that is, whether they *parallel,* thus revealing positive correlation (Figure 112 below), or *mirror,* thus revealing negative correlation (Figure 113 below).

The necessary comparison of the two curves in this method of determining correlation may be facilitated in different ways. The preferable method in the arithmetic type of curves is to superimpose one on the other through "light analysis" or by drawing one on a transparent medium such as tracing cloth or thin paper. If the curves tend to *mirror* each other one of them should be turned upside down or the scale for plotting it should be reversed before they are brought together. This superimposition will readily reveal any correspondence or similarity in the movement



(From article by **M. C. W. Tomlinson**, *Management Engineering*, Dec. 1921)

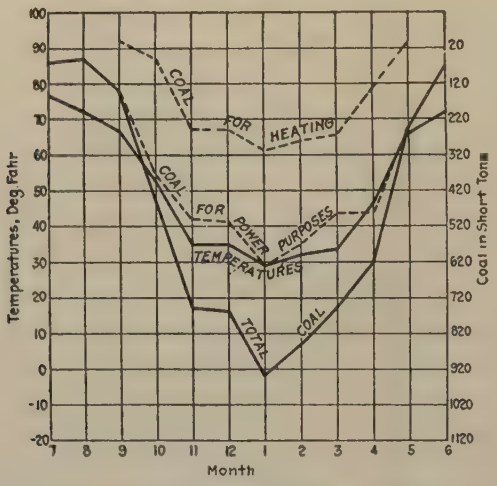
FIGURE 112.—Water and Coal Consumption Curves of a Large Heating Station

Curves that parallel revealing positive correlation.

of the curves as shown in Figures 112 and 113. Since this aspect of the method involves largely the question of bringing two curves together for comparison, the logarithmic graph will be very serviceable.

This method does not yield the exact measure of correlation that is obtained through the computation of a coefficient of correlation, nevertheless because of its simplicity and ease of application it is extremely serviceable in practical business statistics and is used perhaps more extensively than any other method. It af-

fords a quick, clear picture of the relationship between the curves which is generally sufficiently precise for practical purposes. It also enables us readily to detect a "lag" in the correlation of the two series which can easily be taken care of in superimposing the curves. This recognition of the "lag" is particularly serviceable in certain problems of forecasting in practical business. As noted on page 212 of King's "Elements of Statistical Method," we must take great care not to assume causation from the relationship of the curves. This must be established in other ways and must be kept in mind in determining which of the curves *lags* behind.



(From article by M. C. W. Tomlinson, *Management Engineering*, Dec. 1921)

FIGURE 113.—Curves Showing Variations in Coal Consumption with Average Temperature for the Period January, 1911 to January, 1912

Curves that mirror revealing negative correlation. Note the inverted scale.

More Exact Measurements.—The most exact measure of correlation is determined through the computation of a coefficient of correlation. Such a coefficient may be computed in a number of different ways. Those usually considered in the less advanced works on statistical methods are (1) the method of concurrent deviations and (2) the Pearsonian or sum-product method. Since

these devices have little place in elementary practical statistical work, their discussion is omitted from the present treatise. The latter is briefly considered in Appendix E.

It should be noted also that the coefficient is usually computed from sample data and that the coefficient obtained from different samples will probably vary in value due to the effect of sampling alone. These values will also vary from the true value which would be obtained by studying the entire "universe" and the measure of these probable variations is determined through the computation of the "probable error" which is discussed in the references given at the close of this chapter.

Practical Service from Correlation.—Great service can be derived from a knowledge of the relationship between series of data. Rough measurements of this relationship can be determined satisfactorily: it is in the effort to express it with too great a mathematical preciseness that we are confronted with definite limitations. But this mathematical exactness is not necessary for practical purposes; the rougher measures are adequate and in their use we may derive valuable service from correlation.

Is it too much to expect, however, that adequate effort may produce correlation between different series in practical business statistics with such definiteness that by knowing the behavior of the one we may determine directly therefrom the general behavior of the other? Although it is generally true at the present stage of development that in business statistics we can rarely establish definite mathematical relationship between two series, it is equally true that until recent years no great effort has been put forth by business management to determine such relationships and that through the rougher measures, particularly the graphic method, we can in certain instances arrive at conclusions sufficiently exact to be of great service.

In the same way, correlation is serviceable in certain cases of forecasting. When there is a "lag" in the correlation between two series, by inspecting the one series we can extrapolate the general course of the other for a future period equal to the "lag."

Finally, it should be emphasized that the foregoing is but a brief and very limited treatment and that elaborate developments

in the field of correlation should not be attempted without more extensive inquiry into the subject. Other methods of developing measures of correlation are discussed in the references below.

References

Davies, George R. Economic Statistics, 1922, Ch. VI.
Forsyth, C. H. Mathematical Analysis of Statistics, 1924, Ch. X.
Jerome, Harry. Statistical Method, 1924, Ch. XV.
King, Willford I. Elements of Statistical Method, 1912, Ch. XVII.
Mills, Frederick C. Statistical Methods, 1924, Ch. X (Linear), Ch. XI (Time Series), Ch. XII (Non-linear).
Rietz, H. L. Handbook of Mathematical Statistics, 1924, Ch. VIII.
Rugg, H. O. Statistical Methods Applied to Education, 1917, Ch. IX.
Secrist, Horace. An Introduction to Statistical Methods, 1917, Ch. XII.
West, C. J. Introduction to Mathematical Statistics, 1917, Chs. VII, IX.
Yule, G. U. An Introduction to the Theory of Statistics, 1919, Ch. IX.

CHAPTER XLIII

THE TIME SERIES

Fundamental Elements.—Comparison of business facts in different periods of time is one of the most effective and extensively used methods of analysis in the management of a business organization. It is common practice to prepare statistical statements and graphs to bring out a comparison of months, quarters, or years. Recognition of the seasonal character of business data is evidenced by the fact that statistical statements commonly provide a comparison of the current month, not only with the "last month" but with the "same month last year." Thus, comparisons as between different time periods are made of revenues, expenses, production, sales, profits, and all important classes of data. The value of such comparisons is clearly recognized and freely used in the management of all sizes and types of business concerns.

In recent years efforts have been made to develop statistical methods which would reveal fundamental tendencies in series of data over a period of time. These developments have recognized four primary groups of elements which enter into time variations.

First, the long-time or secular trend which represents the general underlying course or direction of the series primarily as upward or downward.

Second, short-time fluctuations which represent the more or less regular periodic movements in the series such as the variation in telephone calls during the 24-hour daily period, and the yearly variations in the residence consumption of electric power, building contracts awarded, and many other classes of data. The short-time fluctuations of most general imoprtance are those recurring from year to year known as *seasonal variations.*

Third, long-term fluctuations known as the cyclical movement or in its general form, the business cycle. While this and the second group are alike periodic fluctuations, those which recur yearly or seasonally and those which recur at intervals longer than a year are distinguished by the terms "seasonal" and "cyclical" fluctuations.

Fourth, irregular or accidental factors such as the effect of wars, floods, pestilence, and other similar phenomena. These cannot be so readily controlled in the analysis but over a long period their influence is relatively not so important as seriously to hinder the more permanently fundamental factors noted in the other three groups.

This chapter will be devoted to an elementary consideration of the broader phases of the methods commonly used to isolate and evaluate the first three types of factors.

It should be noted in this connection that in undertaking to determine the secular trend and typical seasonal variations great care must be given to the period chosen for analysis. Other things being equal, the longer the period the more satisfactory will be the results obtained, particularly for projection beyond the period covered by the data. In our desire to get a long period we should take care not to select data which are obviously unfitted for the purpose: particularly should we avoid a period which includes two or more obviously unlike trends or periods in which accidental influences have been unusually great. This point is discussed more fully in the references listed on page 483.

THE SECULAR TREND

Free-Hand Method.—The simplest method of indicating the secular trend is to plot the actual data as a straight-line curve and then throw a trend line across it free-hand without any computations. The position of such a trend line will not be so mathematically precise as that of others. It depends primarily on the judgment of the statistician and consequently will be constructed differently by different persons. However, in many cases in practical work this rough approximation is sufficiently accurate for the purpose at hand. It is particularly serviceable where the time available for constructing the line of trend is short.

Method of Simple Averages.—The computation of simple averages will assist in the construction of the trend line. The average for the entire series may be computed and plotted on the middle ordinate, the line of trend being drawn through this point. This aids the eye but still leaves a great deal to the judgment.

The series may be divided into equal time groups and the average of each group plotted on the middle ordinates of the two groups. This method was used in the construction of the "semi-average" curve in Figure 114 below. The average for the first half of the series—*i.e.*, the years 1919 and 1920 and the next six months of 1921—is plotted on the ordinate half-way between March and April, 1920, and the average for the second half between September and October, 1922. The line of trend is drawn through these two points. This method is called the method of *semi-averages.*

Where the general movement of the series is irregular, resulting in a broken trend line, the series may be divided into more than two groups and the averages for the separate groups computed and plotted as indicated in the case of the semi-averages. This method is generally used only where there are definite and abrupt changes in the direction of the trend.

The Moving Average.—The method of moving averages previously discussed gives a more precise mathematical character to the trend line than either of the methods noted above. In the use of this method the average of the values for a selected number of time periods is computed and taken as the trend value for the middle of the time period represented. Thus, in the 12-month moving averages in column "c" of Table 1 below, the figure 472.4 is the average of the twelve months of 1919 and is placed at the middle of the year—*i. e.*, between June and July; 483.5 is the average of the twelve months beginning with February 1919 and ending with January 1920 and is placed between July and August. The other items in column "c" are determined in similar manner.

Since it is important in further analysis that the moving averages coincide with the time designations, when the number of time periods in the moving average group is even an adjustment is made by taking a two-period moving average of the moving averages. The moving averages in column "c" of Table 1 have been adjusted in this manner and are shown in column "d." The figure 477.9 is the average of 472.4 and 483.5; the figure 487.9 is the average of 483.5 and 492.3; and so on. The figures in column "d" are represented by the "moving average" in Figure 114 below.

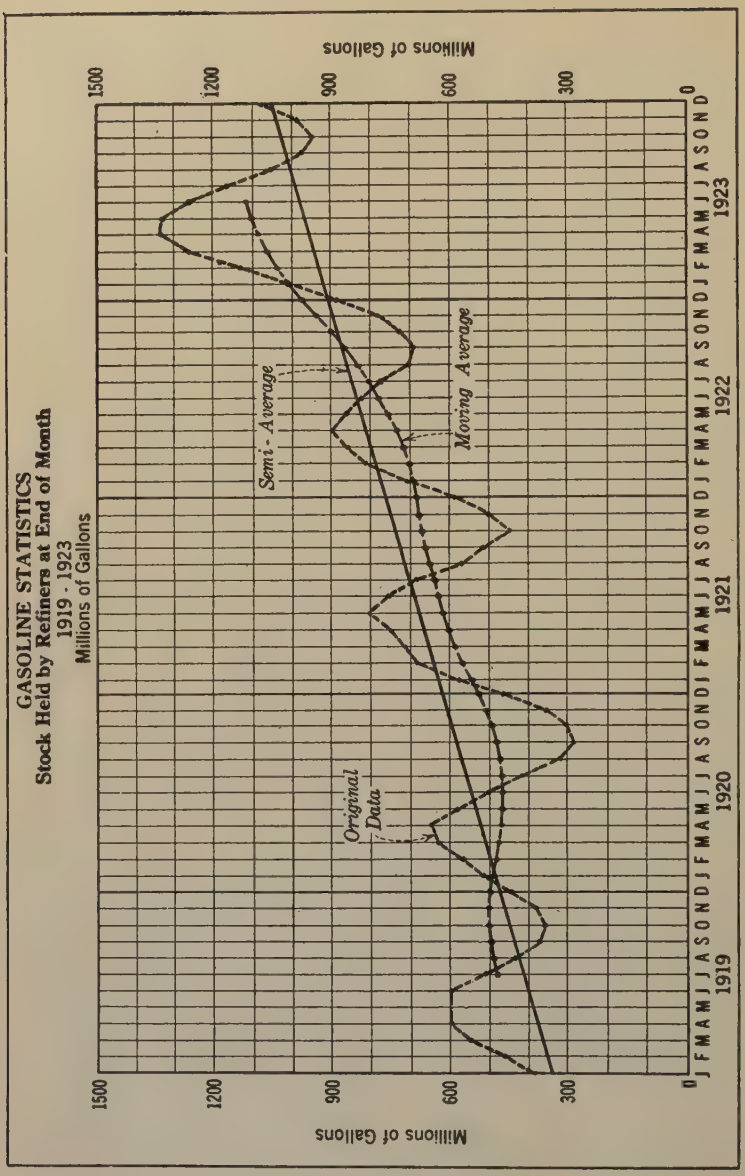

GASOLINE STATISTICS
Stock Held by Refiners at End of Month
1919 - 1923
Millions of Gallons
Millions of Gallons
1500
1200
900
600
300
0
Semi-Average
Moving Average
Original Data
J F M A M J J A S O N D J F M A M J J A S O N D J F M A M J J A S O N D J F M A M J J A S O N D
1919
1920
1921
1922
1923

FIGURE 114.—The Trend Line in a Time Series by the Methods of Semi-Averages and Moving Averages

The important element in the use of the moving average is the selection of the length of the period. The regularity of the cycles in the original data, the regularity of the amplitudes of the fluctuations, and the linearity of the trend are primary factors in

GASOLINE STATISTICS.
Stocks on Hand by Refiners at End of Month
1919-1923.
(In millions of gallons.)

Years and Months	Original Data	12 Months Moving Average	2 Months Moving Average of Column c	Years and Months	Original Data	12 Months Moving Average	2 Months Moving Average of Column c
a	b	c	d	a	b	c	d
1919—J	383			1922—J	706	694.7	691.0
F	458			F	807	706.0	700.4
M	546			M	854	720.6	713.3
A	594			A	893	742.9	731.8
M	594			M	857	766.3	754.6
J	594	472.4		J	825	791.1	778.7
J	515	483.5	477.9	J	773	815.9	803.5
A	435	492.3	487.9	A	704	842.8	829.4
S	371	498.9	495.6	S	690	876.6	859.7
O	354	503.1	501.0	O	724	913.5	895.1
N	378	501.8	502.5	N	777	952.8	933.2
D	447	494.3	498.1	D	884	989.4	971.1
1920—J	516	485.8	490.1	1923—J	1003	1022.1	1005.8
F	563	476.4	481.1	F	1130	1051.3	1036.7
M	626	469.5	473.0	M	1259	1074.8	1063.2
A	644	465.1	467.3	A	1336	1093.4	1084.1
M	578	463.2	464.2	M	1329	1110.7	1102.2
J	504	464.4	463.8	J	1264	1126.6	1118.7
J	413	469.1	466.8	J	1165		
A	323	478.9	474.0	A	1054		
S	288	486.1	482.5	S	973		
O	301	494.7	490.4	O	947		
N	355	513.3	504.0	N	985		
D	462	533.9	523.6	D	1075		
1921—J	572	556.5	545.2				
F	681	576.9	566.7				
M	713	595.8	586.4				
A	747	608.7	602.3				
M	801	620.5	614.6				
J	751	630.8	625.7				
J	684	642.0	636.4				
A	568	652.5	647.3				
S	515	664.3	658.4				
O	456	676.4	670.4				
N	496	681.0	678.7				
D	586	687.3	684.2				

(Source: Standard Statistics Co. and U. S. Department of Interior)

TABLE I.—Representing the Trend by the Method of Moving Averages

this selection. If the cycles and amplitudes are uniform and the trend linear in character, the selection of the moving average period which will give a good representation of the trend is comparatively simple. But in practical work the cycles and ampli-

tudes are generally irregular and frequently the general trend is non-linear and hence the problem is often difficult. As a matter of fact, in this problem as in all problems in practical work, the decision can be made only after careful consideration of the actual data to be analyzed. However, certain helpful considerations can be noted.[1]

1. In general, the moving average should be based on a period equal to the period of the cycle to be removed or to some multiple of the period of the cycle.
2. If the cycle is irregular, the moving-average period should be equal to the average length of cycle or preferably to a higher multiple of this average. The same is true when the amplitudes vary, the cycles being uniform, or when both the cycles and amplitudes vary.
3. When the underlying trend of the series changes direction the moving-average period should be reduced to a minimum consistent with the length of the cycle.

In practical work the selection of the moving-average period is somewhat of a trial and error process. The data are usually plotted in a straight line curve from which the cycles are apparent and in some instances the period can be readily determined by inspection from this curve. Thus, in Figure 114 it is clear that a 12-month period will be satisfactory, as the cycles are somewhat regularly twelve months in length. Where the cycles are irregular we may arbitrarily take the average or preferably we may try out periods of different lengths and compare the resulting trend lines.[2]

The method of moving averages is very serviceable in practical work primarily because of its simplicity. While it does not represent the trend with the mathematical preciseness that is obtained by the use of some other methods, it is usually sufficiently accurate for practical purposes, particularly when the length of the cycle is approximately regular and the underlying trend is linear. It cannot, of course, be extended the full length of the period under

[1] Further discussion will be found in the references listed on page 483, particularly on pages 262 to 271 of Professor Mills' "Statistical Methods."

[2] Weights are sometimes applied to the data in computing the moving average, the items near the point at which the average is centered being given the heaviest weights. Binomial coefficients are often used, 1, 2, 1 for average of three items, 1, 3, 3, 1 for four, 1, 4, 6, 4, 1 for five, and so forth. This form of the moving average is called the progressive mean.

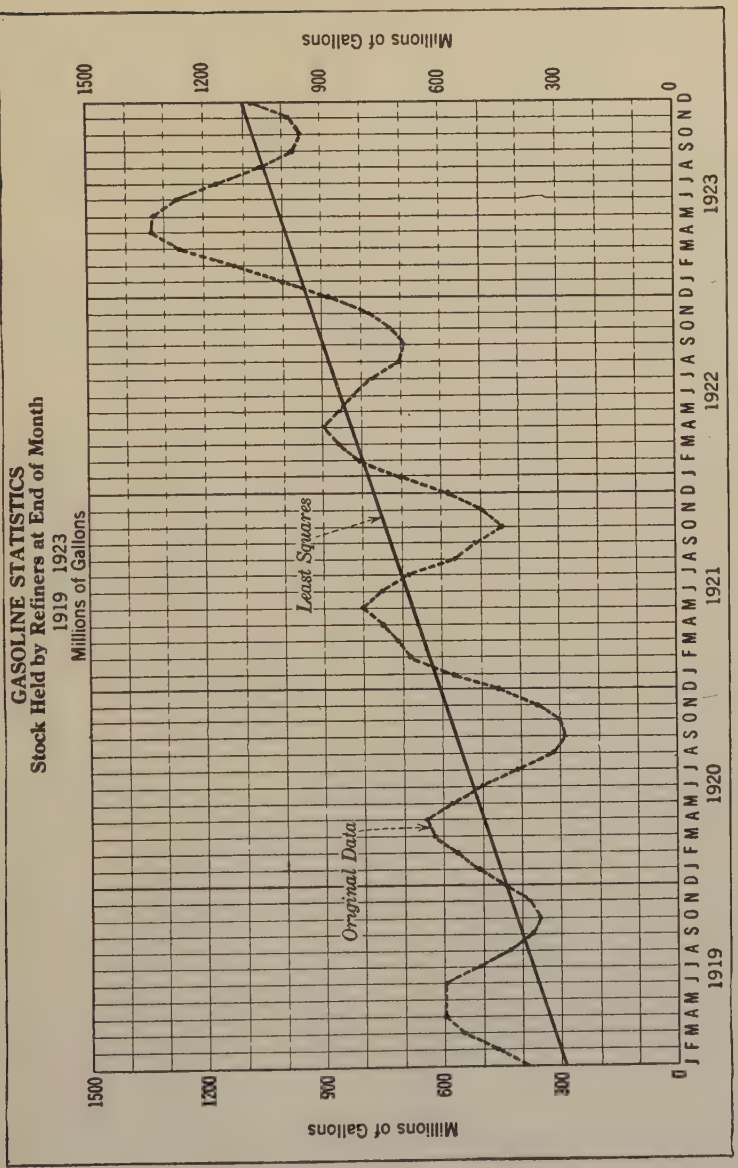

GASOLINE STATISTICS
Stock Held by Refiners at End of Month
1919 1923
Millions of Gallons
Least Squares
Original Data
Millions of Gallons
1500
1200
900
600
300
0
J F M A M J J A S O N D J F M A M J J A S O N D J F M A M J J A S O N D J F M A M J J A S O N D
1919
1920
1921
1922
1923

FIGURE 115.—The Trend Line in a Time Series by the Method of Least Squares
For the ordinates of the trend line and the method of determining them, see Appendix E, Part 3

consideration without interpolation at either end for a period equal to one-half of the moving-average period.

Method of Least Squares.—The method of least squares gives a more truly representative line of trend than any of the other methods. Detailed development of this method is given on page 585 of Appendix E and the trend line of the data of Table 1 above determined by this method is shown in Figure 115 above.

The chief objections to this method in elementary work are the somewhat complex mathematical processes and the extensive computations involved in its use. However, where accurate results are required it is preferable in most instances.

SEASONAL VARIATION

Method of Simple Averages.—There are three principal methods of analyzing the seasonal variations of a time series—(1) simple averages, (2) moving averages, and (3) link relatives.

The first of these consists fundamentally of taking simple arithmetic averages of the items separately for the recurring time periods. In Table 2, for example, the items for each month over the five-year period are averaged. Thus, 636.0 in column "g" is the arithmetic average of 383, 516, 572, 706, and 1003, the five January figures. The averages for the other months are determined in similar manner. These monthly averages are then converted to relatives with the average of the monthly averages as a base to get a final seasonal index. For gasoline stocks this index is given in column "h."

It is apparent that the secular trend may seriously affect the averages of column "g" and where a high degree of accuracy is required they should be corrected for the secular trend. This is accomplished by determining the average monthly increase in stocks over the five-year period preferably by the method of least squares and, taking some period as a starting point, by adjusting the averages to eliminate the influence of this increment. Thus, if June were taken as the starting point, the monthly increment would be added to the average for May, twice the increment would be added to April, three times to March, and so on to January. On the other hand, the increment would be subtracted from

July, twice the increment would be subtracted from August, three times from September, and so on to December. More detailed description of this method is given in the references listed on page 483.

GASOLINE STATISTICS.
Stocks on Hand, End of Month
1919-1923
(In millions of gallons.)

Months	1919	1920	1921	1922	1923	Average 1919–1923	Relatives of Items in Column g with Their Average as Base
a	b	c	d	e	f	g	h
January	383	516	572	706	1003	636.0	91.2
February	458	563	681	807	1130	727.8	104.4
March	546	621	713	854	1259	800.0	114.8
April	594	644	747	893	1336	842.8	120.9
May	594	578	801	857	1320	831.8	119.3
June	594	504	751	825	1234	787.6	113.0
July	515	413	684	773	1165	710.0	101.9
August	435	323	568	704	1054	616.8	88.5
September	371	238	515	690	973	567.4	81.4
October	354	301	456	724	947	556.4	79.8
November	378	355	496	777	985	598.2	85.8
December	447	462	586	884	1075	690.8	99.1
						697.1	

TABLE 2.—Seasonal Variation by the Method of Simple Averages
The seasonal index unadjusted for secular trend is given in column "h."

Method of Moving Averages.—The moving-average method of determining seasonal variation consists of four steps. These will be illustrated by the application of the method to the data on gasoline stocks as given in Table 3:

1. The moving averages are taken (column "d," Table 1 above).
2. The actual data are expressed as percentages, item by item, of the moving averages (columns "b" to "f," Table 3 below).
3. The percentages determined in No. 2 are averaged month by month for the five-year period (column "g" in Table 3 below).
4. These averages are converted to relatives with the average of these monthly averages as base (column "h" in Table 3 below).

The arithmetic average or the median may be used in computing these averages. The former has been used throughout in

Table 3. For elementary practical work it is generally preferable to use the arithmetic average wherever possible and unless extreme items are present to distort this average it should be used in such computations.

Method of Link Relatives.—This method consists of (1) developing a series of chain relatives through the computation of median link relatives, (2) eliminating the influence of the secular trend from these chain relatives, and (3) arriving at a final index by converting the adjusted series from a monthly to a yearly base. It is somewhat more complicated than we should use in

GASOLINE STATISTICS.
Stocks on Hand, End of Month
1919-1923
(In millions of gallons.)

Months	1919	1920	1921	1922	1923	Average	Seasonal Index
a	b	c	d	e	f	g	h
January		105.3	104.9	102.2	99.7	103.0	103.6
February		117.0	120.2	115.2	109.0	115.4	116.1
March		132.3	121.6	119.7	118.4	123.0	123.7
April		137.8	124.0	122.0	123.2	126.8	127.6
May		124.5	130.3	115.1	120.6	122.6	123.3
June		108.7	120.0	105.9	113.0	111.9	112.6
July	107.8	88.5	107.5	96.2		100.0	100.6
August	89.2	68.1	87.7	84.9		82.5	83.0
September	74.9	59.7	78.2	80.3		73.3	73.7
October	70.7	61.4	68.0	80.9		70.3	70.7
November	75.2	70.4	73.1	83.3		75.5	76.0
December	89.7	88.2	85.6	91.0		88.6	89.1

TABLE 3.—Seasonal Index by Method of Moving Averages

Columns "b" to "f" contain the percentages resulting from dividing the actual data, item by item, by the moving averages. Column "g" contains the averages month by month of these percentages. The seasonal index is given in column "h," which consists of the items in column "g" expressed as relatives of their average,—*i.e.*, each item is divided by the average 99.4

elementary work and consequently the detailed description of the method has been transferred to Appendix E.

Graphic Methods.—A rough test of seasonal variation may be made by plotting the original data in a straight line curve or a column diagram. The "original" curve in Figure 114, for example, suggests the presence of a very definite seasonal variation. If curves representing the separate years were drawn to a single

twelve-months scale the seasonal variation would be brought out even more clearly. The semi-logarithmic grid is most effective for this purpose.

The same condition would be revealed if column graphs were used to represent the plotted points in Figure 114. Or if all the January figures were plotted in succession and then all the February figures, and so on, the seasonal variation would be better revealed.

CYCLICAL MOVEMENTS

Nature of the Problem.—It will be recalled that the terms cyclical movement and business cycle are used to designate the type of periodic fluctuations whose length is more than one year, the shorter fluctuations being known specifically as "daily" and "seasonal" variations. The method of procedure in the analysis of this factor of the time series consists primarily of eliminating from the original data the influence of the secular trend and seasonal variation. The processes are somewhat more complex than those required in the study of the secular trend and the seasonal variation discussed above and a comprehensive treatment of them would fall without the scope of this work. Hence, we shall merely suggest in broad outline some of the methods being used at present to determine the cyclical movement. The movement sought may be the fundamental periodic movement underlying the general economic structure or it may be the behavior of a certain special type of data within the individual business concern.

Eliminating the Secular Trend.—If annual data are being dealt with the problem will be confined to the elimination of the secular trend from the original data. Two principal methods are employed:

1. The yearly figures which represent the trend line, determined preferably by the method of least squares, are subtracted item by item from the original data. This method is fully developed in the references listed on page 483.
2. The original data may be divided, item by item, by the figures which represent the trend line, determined preferably by the method of least squares. Similarly, the deviations of the

original data from the trend figures may be divided by the trend figures. This process is also developed in the references given below.

Eliminating Both Secular Trend and Seasonal Variation.—The elimination of both the secular trend and seasonal variation may be accomplished in different ways. A simple procedure involves the following steps:[3]

1. The ordinates of secular trend and the seasonal indices are determined as outlined above.
2. The actual data are expressed as percentages, item by item, of the ordinates of the secular tread—*i.e.*, each actual item is divided by its corresponding trend ordinate. This process eliminates the influence of the secular trend.
3. The seasonal indices are subtracted from the percentages determined in No. 2. The resulting series represents the deviations from the normal.

These are the principal steps to be followed in determining the cyclical movements of single series of data or their variation from normal such as we usually require in the management of the individual concern. These deviations are reduced to percentages of their standard deviation when we wish to make comparison between different series. Further steps in the development of a general business cycle have to do with methods of combining cycles for different series of data and are too complex for a treatise of this kind. Extensive discussion of this entire method of analysis may be found in the references given on the following page.

General Considerations.—The foregoing brief discussion of the more elementary aspects of the application of these methods of statistical analysis is designed merely to serve as introduction to this general type of method which has developed so rapidly within recent years rather than to represent a survey of such development. A complete and comprehensive review of the various methods that have been developed for analyzing the time series could in itself well occupy a volume of considerable size. In view of the facts that these methods are not yet standardized and that

[3] G. R. Davies, "Economic Statistics," page 120.

their successful application depends so much on the character of the data, we should not attempt to use the more accurate methods without first thoroughly mastering the method and applying it in the analysis of practice problems. In practical elementary work we should confine our use of such methods to the rougher approximations and should interpret the results only as rough approximations.

References

Berridge, W. A. Cycles of Unemployment in the U. S.—1903-1923, (1923).
Davies, George R. Economic Statistics, 1922, Ch. V.
Jerome, Harry. Statistical Method, 1924, Ch. XIII (Secular Trend and Seasonal Index), Ch. XIV (Cycles and Barometers).
Karsten, Karl G. Charts and Graphs, 1923, Ch. XXI.
King, Willford I. Elements of Statistical Method, 1912, pp. 167-178.
Mills, Frederick C. Statistical Methods, 1924, Ch. VII (Secular Trend), Ch. VIII (Seasonal and Cyclical Fluctuations).
Mitchell, Wesley C. Business Cycles, 1913.
Moore, H. L. Economic Cycles: Their Law and Cause, 1914.
—— Generating Economic Cycles, 1923.
National Bureau of Economic Research. Business Cycle.
—— Unemployment, Employment Hours and Earnings in Prosperity and Depression. United States 1920-23, (1923).
Persons, W. M. Indices of Business Conditions. The Review of Economic Statistics, Preliminary Vol. I, 1919, pp. 5-107.
Rietz, H. L. Handbook of Mathematical Statistics, 1924, Ch. X. (Warren M. Persons).

CHAPTER XLIV

NATURE, IMPORTANCE AND METHOD OF INTERPRETATION

Meaning of Interpretation.—Following analysis, the next step in the development of a program of statistics is interpretation. While it is difficult to formulate specific directions for this step, we shall undertake in this chapter to indicate what interpretation is, to offer practical suggestions relative to methods of interpretation, and to point out the importance of interpretation in practical statistical work.

It is generally assumed that the meaning of interpretation is clear, and consequently it is usually not included as a separate subject in a series of discussions of this kind. While this assumption may be true in its general aspect, an inspection of the interpretative comments on statistical analysis in practical business statistics will lead to serious question of its correctness. The conclusion from such observation is that interpretation in practical work is confused with analysis and as a result an extension of analysis is frequently offered as interpretation and real interpretation is neglected altogether. Hence, our effort to define interpretation shall take the form primarily of distinguishing it from analysis.

It was pointed out on page 354, that a primary aim of analysis is to facilitate interpretation. More specifically analysis, through the various means discussed in Chapters XXXIII-XLIII, determines from the great mass of unsystematized detail what the important facts and conclusions are. Interpretation begins where analysis leaves off. It is concerned solely with the determination of the meaning or significance of the facts or conclusions which have been established through analytical processes.

For example, a so-called interpretative memorandum may state merely that the average rate of pay of a certain class of employes is a certain percent higher in one department than in another.

But this is not interpretation: it is analysis; and this confusion is serious because no attempt is made really to point out the meaning of the fact in its relation to problems of management. What is the effect of the disparity in rates on the cost of production, the amount of working capital required, labor turnover, quality of employes, and so forth? Does it mean that one department is managed more efficiently than the other, and, if so, in what specific respect? Does the one department keep a higher grade of help, maintain more efficient supervision, have better employment conditions, and hence longer average period of service? These and many similar questions will be raised in our effort to evaluate the fact.

In brief, interpretation will involve thorough consideration of two general aspects—why are the rates different, and, in the light of this explanation, what is the significance of the fact in the management of the business. The explanation of the disparity in rates may be found in a number of related facts such as the difference in period of service, grade of employes, rate schedule, and so forth. The meaning of the fact will vary with the explanation. If, for instance, the disparity is due to difference in period of service, the department with the higher rate of pay may, contrary to first impressions, be more efficient since the longer period of service means more experienced employes and hence may result in a smaller number and a relatively smaller payroll outlay. Both of these aspects of the situation—the *explanation* and the *meaning* of the fact—must be fully and thoroughly developed in arriving at a final interpretation of the fact.

Interpretation, therefore, may be defined as the process of determining the meaning or significance of facts or conclusions in the management of the business and consists of establishing both the *explanation* and the *meaning* of the data. The development of related facts required in the explanation is, of course, an analytical process but the determination of their bearing on the fact to be interpreted is a part of the interpretative process. The importance of maintaining this conception in practical work, particularly as distinguished from analysis, cannot be overemphasized. More is said on this point in the discussion of the importance of interpretation on page 489.

Aim of Interpretation.—The aim of interpretation is to translate the facts finally revealed by analysis into bases for action. This is the step which the statistical service frequently fails to take, a fact which in many cases seriously cripples the service. Thus, a statistical analysis may compare the rates of pay of machinists in the company with corresponding rates paid generally by outside concerns in the community and may establish the fact that the company's rates are lower than those paid by other establishments. Merely to establish this fact, however, is of little importance to anyone in the company. The real service of this fact comes from the determination of its meaning; we want to see why our rates are lower and the effect the fact is having and may have on the business.

Such a fact may mean any one or a combination of a number of things. It may mean that the company's rate is too low and that there is danger that the company may suffer therefrom in labor losses, discontented employes, inability to get good grade of workers, and so forth; it may mean that the company does not require employes so highly skilled and experienced as other concerns and that after all the disparity reflects only proper differences; it may mean that the difference in pay is satisfactorily offset by shorter hours, steadier employment, better employment conditions in the form particularly of sickness, accident and death benefits, restaurant service, savings plan, stock-purchase plan, and so forth, which enable the company to compete favorably in the market in spite of the disparity in rates of pay.

Now, it is this latter function—the determination of the meaning of the facts—which is the primary aim of interpretation. It evaluates fully and completely the important facts established and presents them in such way that in the hands of the management they become at once the basis for action.

Methods of Interpretation.—The method of procedure in interpreting a statistical statement consists primarily of three broad steps—(1) familiarizing ourselves with the problems of management in the particular field to which the data relate, (2) noting the important facts and conclusions brought out by the study, and (3) determining specifically the significance of these facts and

conclusions as related to the management of the business—*i. e.*, the bearing of the facts of (2) on the problems of (1).

Review the Principles and Problems of Management.—This step consists primarily of reviewing the principles of management involved in the particular phase of the business to which the data relate and calling clearly to mind practical problems involved in their application. If the facts are in the personnel field, we must be familiar with both the principles and practices of personnel management. The same is true for financial or operations facts. It is the height of folly to assume that a statistician can effectively and wisely interpret business facts without being familiar with the problems of management to which they relate. He may be able to analyze the data and even to note the outstanding conclusions without such knowledge but his interpretations cannot be depended on unless he renders them with intimate knowledge of the related problems of management. As a matter of fact, the analytical program can be much more effectively directed if the person in charge is also familiar with the principles and problems of management.

This is one of the most important steps in the interpretation of statistical data and is the reason why in so many cases the interpretation of the facts is left entirely to the executives. It is assumed that the executive is the only person sufficiently familiar with the management problems to be able properly and effectively to make the interpretations and in far too many cases this assumption is based on fact.

It is because of this fact that a practical knowledge of business management is emphasized on page 532 as a qualification of the executive and staff statisticians. The statistician may be able to compute all sorts of averages, work up perfect index numbers, develop a thoroughly satisfactory frequency classification, and do the other forms of analysis well, and still not be qualified to interpret business facts simply because he is not practically familiar with business management. In such case it is generally better that he make no effort to do so and leave it entirely up to the executive who does know what is important in directing the affairs of the business. This is in no small part the reason why the business

man places so much stress on practical experience in the selection of his personnel.

Important Conclusions.—The second step in interpretation is to note the important facts and immediate conclusions brought out by the statistical analysis. This is really a phase of analysis and might well have been included in the discussion of that subject but because of its intimate association with the process of interpretation it has been introduced at this point.

Perhaps the most effective and most extensively used method of determining the important facts brought out by the analysis is comparison. We may compare (1) the facts at one time with corresponding facts at another time, say, this year with last year or this month with last month or the same month last year, (2) facts of one organizational unit with corresponding facts of another such as one department, division, branch or company with another, (3) actual facts of the business with requirements or with carefully determined *standard facts* or with arbitrarily established "bogies," (4) individual items or summations with totals of which they are a part, (5) the facts of the business with certain other facts, within or without the business, to which they are supposed to bear a fixed relationship, (6) the facts of the business with other facts or conditions which are supposed to influence them.

This is not the only method, however. An inspection of the facts independently may be all that is necessary. Thus, the deliveries to be made indicates the transportation facilities to be provided, the average rate of pay and the number of employes indicates approximately the funds to be provided for payrolls, and so forth. In the great bulk of statistical analyses, nevertheless, comparison is used to bring out the facts of importance. This service is next in importance to the actual interpretation of the facts, and, as stated above, the statistician can pretty well indicate the important conclusions resulting from analysis even though he is not intimately familiar with the problems of management. The mere indication of the important conclusions is very serviceable to the executive even though no attempt is made at interpretation. It relieves the executive of the necessity of wading through the tables of figures

or of spending his time examining graphic presentations. If the conclusions are at once apparent to him, he can with little delay turn them over to his department heads or staff assistants for interpretation and recommendations as to executive action. There is no excuse for a failure to point out what the important conclusions are in a statistical study or report. The statistical organization which as a regular practice neglects this important step is offering a limited and ineffective service.

It should be noted, however, in this connection that great care should be taken in accepting conclusions as finally established. So much of our work in the field of practical business statistics is based on estimates and final results are so frequently mere indications, that we must be extremely cautious to take the conclusion for just what it is worth.

Determining What the Conclusions Mean.—The final step in interpretation for the accomplishment of which the other two are but preparation is the determination of the meaning of the conclusions. It is in fact the only part of the process which is really and exclusively interpretation and is therefore the fundamentally important phase of the process. The detailed method of procedure in this step consists of evaluating the facts and conclusions in the light of a certain purpose.

The two steps noted above will give us a list of the facts and conclusions resulting from the analysis and a list of the general phases and specific problems of management to which they severally relate. The beginner will find it helpful to take the conclusions in order and consider them, one by one, in their relation to each of the phases and problems of management. This procedure will give a systematic and comprehensive evaluation of the facts and conclusions established through the analysis of the data.

The Importance of Interpretation.—As emphasized throughout the chapter, interpretation is the crowning feature of the statistical study. It is the goal toward which each step in the development of the program of statistics leads. The collection of data, their recording, their presentation, and their analysis are worth while only in case they are finally interpreted to throw light on the practical problems in the management of the business. It

is this determination of the significance of the facts and conclusions established through analysis that creates concrete bases for action and thus produces the final rewards for statistical efforts.

Hence, the main emphasis of the statistical service should be placed on interpretation. The statistician is prone to devote too much time to assembling and analyzing business facts and too little time in establishing important numerical conclusions in interpreting them. Even though an admirable system of basic data records is maintained and the facts furnished thereon carefully and systematically analyzed, the statistical service falls short unless the conclusions established by the analysis are clearly indicated and their significance in the management of the business pointed out.

This is not to say that no benefit is to be derived from the type of service which stops with analysis. There are many such services to be found in practical business concerns and much good is being derived from them particularly where the conclusions are noted for the executives. But such a service places the entire burden of interpretation on the executive and it would be much more effective and beneficial were the analyzed data accompanied with a reasonably complete interpretative statement. Such a practice will tend to relieve the executive of the burden of working out interpretations and will furnish the facts finally translated into bases for immediate action. The executive may supplement or modify the interpretations presented by the statistician but these contain the essential features and reduce the work of the executive on this account to a minimum.

The question is frequently raised as to how far the statistician should attempt to go in interpretation. The answer is clear. Generally speaking, if he is qualified he should attempt a full and complete interpretation and, if he is not qualified, he should make little or no attempt along the line. Particularly should he undertake to offer interpretations of conclusions which apply to the company as a whole; those relating specifically to departmental affairs may preferably be handled by department heads.

Distributing Business Facts.—It may be noted at this point that attention can well be given to the proper method of distributing business facts. There are three principal methods of distri-

bution. First, information may be furnished informally as in the case with a multitude of minor inquiries; second, routine reports (informative or interpretative) may be developed which are distributed periodically; and, third, special studies may be worked up as occasion requires.

Information distributed informally is generally handed out in response to specific inquiries and for the most part the data called for are few. The chief benefit of the method lies in the fact that the data may be obtained from the statistical organization in a minimum amount of time.

Distribution by routine reports involves the preparation of formal statements at periodic intervals and their distribution regularly to a specific number of recipients. It is the method most generally employed in practical statistical organizations. However, simply because it is the most effective method of getting out a great volume of data, its use frequently leads to extravagance in time and money and tends to lower the quality of service through placing emphasis on the volume of data that are worked up and distributed rather than on their analysis and interpretation. What has been said is not to be interpreted as an attempt to advise the elimination of routine reports. It is against the wholesale preparation and distribution of such reports and the resultant extravagance that caution is being urged. Routine reports are an important part of a well-rounded program of statistics and as such should receive very careful consideration in their development. They are particularly serviceable in furnishing to the executives regularly certain important facts of the business which are used as broad checks on the progress of the enterprise. It should be emphasized, however, as already pointed out, that economy can be frequently realized in the use of routine reports by substituting a brief memorandum for a particular issue providing no material change has occurred since the last issue.

Special reports both written and oral are of an entirely different character. They are specific replies to definite inquiries. The very existence of these inquiries presupposes the attention of the inquirer on the subject of the report if the inquiry has come from outside the statistical organization. If the special study has been originated in the statistical organization then its newness adds to

its interest and tends to attract the attention of the recipient. In general they fall within the following classes:

1. Studies covering that portion of the basic data which has not been touched upon in the development of routine reports.
2. Studies in the same field as routine reports but going into greater detail. These studies are frequently the outgrowth of inquiries arising from inspection of the routine reports.
3. Research studies involving special developments of certain data from the basic records or through special collection from sources within or without the company.

Part V— Miscellaneous

CHAPTER XLV

MECHANICAL AIDS IN STATISTICAL WORK

Need for Mechanical Aids.—The clerical labor involved in handling great masses of data is often a serious handicap in practical statistical work. It is only through the assistance of various mechanical devices that we are able satisfactorily to accomplish many of the more important statistical developments in business concerns. It is the purpose of this chapter to discuss briefly the broad aspects of the operations of certain of these devices and the services which may be derived from them. The aim is not so much to develop detailed instructions for operating these devices as to outline the services which may be derived from them.

Computing Machines.—The various computing machines used in statistical computations are generally familiar. Some of those commonly used are the Burroughs, Monroe, Comptometer, Ensign, Wales, Dalton, Barrett, Millionaire, Bruns Viga, and so forth, a very complete list of which may be found in "The American Digest of Business Machines," published by The American Exchange Service, Chicago. Most of these machines come equipped for hand or electric operation and are designed to perform the operations of addition, subtraction, multiplication, and division. They are all easily operated and will generally give equally satisfactory service. Some print the results of the operations on tape paper, such as the Burroughs, and thus provide a means of checking the results, while others, such as the Comptometer, make no such records, repetition of the process being the only means of checking the results.

Needless to say these machines are used for practically all computing work in business organizations today. Certainly, for any extensive operations in addition, subtraction, multiplication, or division we should not think of doing the work longhand.

Some of the machines are more serviceable for certain operations and other machines for other operations. The selection of a machine or machines should be made only after careful consideration of the different types in the light of the services to be performed. It should be noted in this connection that expert operators will be preferable if extensive production is required, particularly with machines of the type of the Comptometer.

In this connection also should be noted the computing typewriter, billing machines, and other machines whose primary service is of another type but which may be used in statistical work.

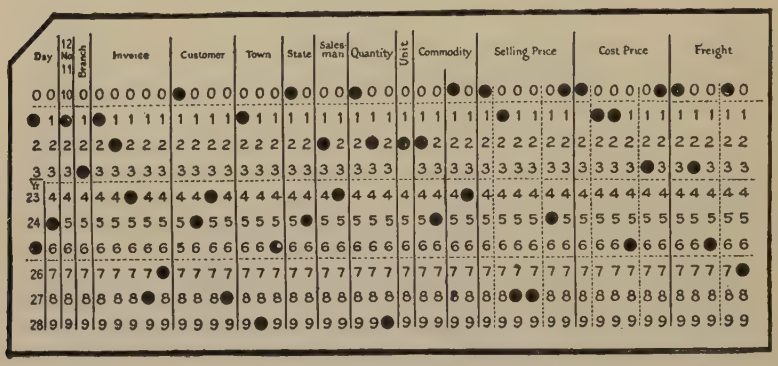


FIGURE 116.—Tabulating Card
Figures 116, 117, 119-123 are used by courtesy of Tabulating Machine Company

The Hollerith System.—The two tabulating machines in most general use are the Hollerith and the Powers. Our discussion will be confined to a consideration of the Hollerith system. This system is typical and the principles of operation and services rendered are not materially different in the different types.

The Hollerith system was developed to facilitate the handling of detailed data in connection with the Federal Census. It consists fundamentally of an electrical device for classifying and tabulating data and its chief service comes from the fact that through its use a great volume of detailed work can be accomplished in a relatively short period of time. The base of the Hollerith system is a specially designed card which is prepared and used by means of the punching, verifying, sorting, and tabulating machines. The methods of preparing the card and of operating the machines are

generally simple and the work can be readily grasped by office employes.

The Hollerith Card.—The Hollerith card contains in its more regular form a series of vertical columns of digits from 0 to 9 and is clipped at one corner to aid the operator to keep the card in the same position for all operations. Figures 116 and 117 are typical cards. It is frequently advantageous to use differently colored cards to represent different groups of data such as, for example, different colors for men and women employes, different

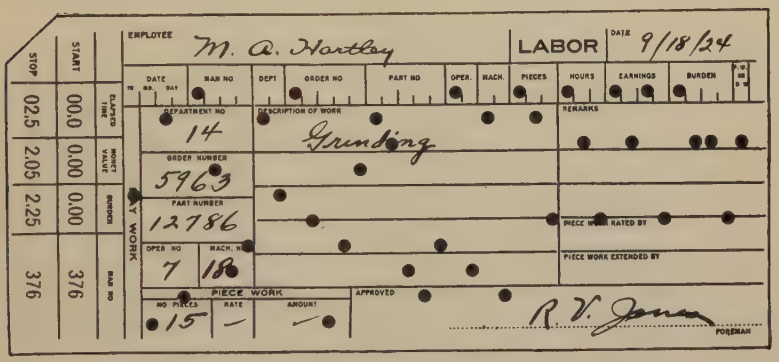


FIGURE 117.—Dual Tabulating Card

sales districts, different commodities, different materials, or what not. In the preparation of the card the following steps are involved:

1. *Select the data that are to go on the card.*—These may be almost any kind of personnel, financial, or operations facts. For example, we may wish to record complete information as to rates of pay, in which case a card would be prepared for each employe, showing perhaps the department, division, occupational class, sex, period of service, and so forth; or we may wish sales data, in which case a card may be prepared for each sale, showing the salesman, territory, price, date, quantity, grade and so forth, to suit the requirements of the particular case.

   We should generally be liberal with the data which are

placed on the cards since this makes them readily available with a minimum expenditure of time and money.

2. *Determine the codes and select the fields.*—In many cases the information will have to be coded before it can be recorded on the card. For example, if we have four divisions in the company we may give them the codes 1, 2, 3, 4. A similar procedure will be followed for departments, occupational classes, salesmen, districts, classes of goods, and other non-numerical facts. Numerical facts such as age, period of service, rates of pay, price, and so forth, can generally be recorded directly but in such cases also codes may be used if economy of space requires it.

   In conjunction with the selection of codes the "fields" of the card will be determined also—*i.e.,* the vertical columns will be grouped and assigned definitely to specific codes. The number of columns to be included in any "field" will be determined by the maximum number of digits in the number to be recorded therein: thus if a code contains two digits it will require a "field" of two columns to record it, or, if a price contains four figures it will require a "field" of four columns to record it, and so forth.

   The chief point of concern in this connection is to economize space as much as possible in order to get a maximum amount of data on a single card. Figure 116 above will illustrate these points.

The Punching Machine.—The data are recorded on the cards simply by punching holes in the card at the points corresponding to the figures to be recorded. In Figure 116 above, for example, a sale price of $0.55 is represented by punching holes at the points 0 in the first 4 columns and 5 in the last two, of the field headed "Selling Price."

These holes are punched by means of small machines operated electrically or by hand similarly to computing machines except that there are much fewer keys. These machines are of three types—(1) the *key punch* (Figure 119 below) which is used for punching individual cards, (2) the *gang punch,* which is used when the same information is to be recorded on several cards such as the day, month, department, division, or what not, in which case several cards (as many as 12) are punched at one time, depending

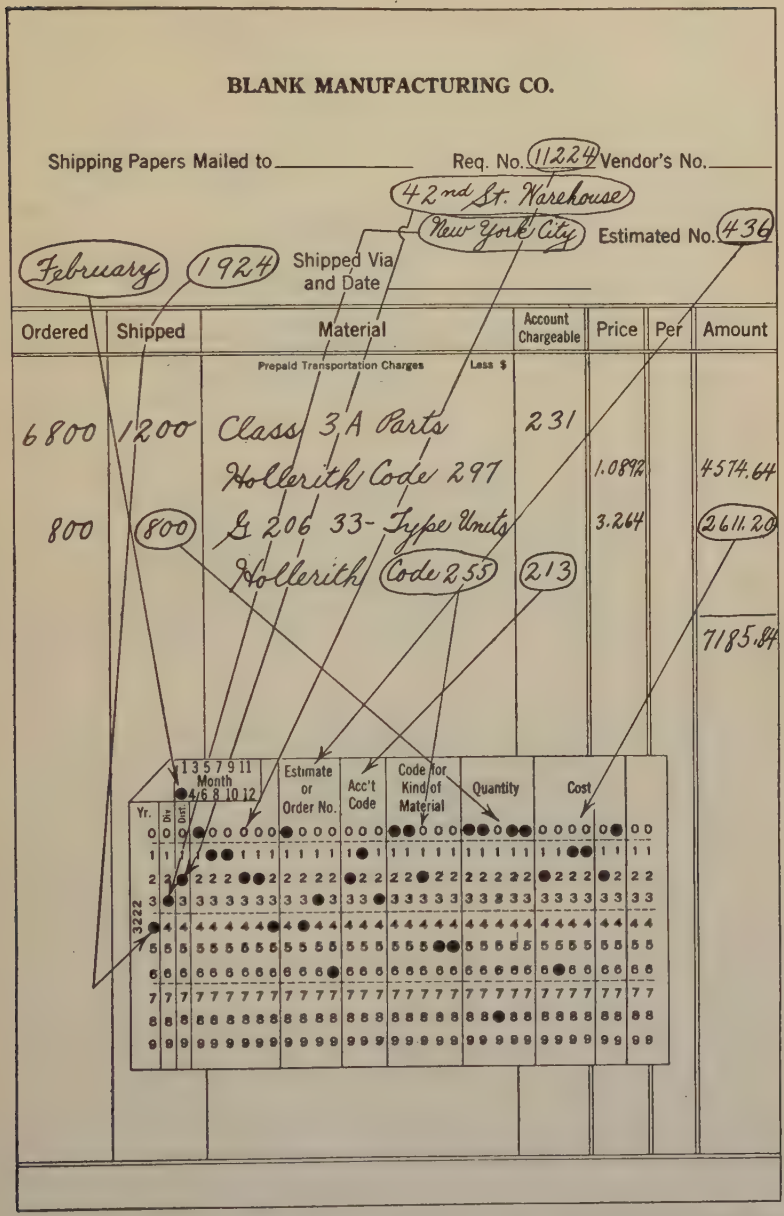


FIGURE 118.—Illustrating the Method of Representing Data on the Tabulating Card

on the number of perforations to be made, and (3) the *duplicator machine* described in the next section.

Experienced operators become very skillful at these machines and can accomplish a great volume of work depending on the character and amount of the information to be recorded.

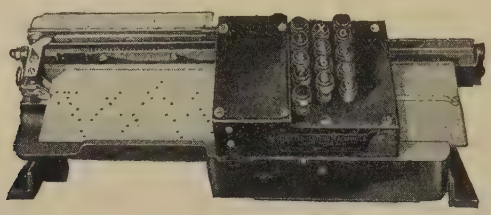


FIGURE 119.—Electric Key Punch

Electric Key Punch with Duplicating Attachment.—The electric key punch with duplicating attachment (Figure 120 below) is an improved method of accomplishing the same general service as the gang punch. It accelerates the recording of data on tabulating cards by automatically punching all the information common to a number of cards. This is accomplished by first

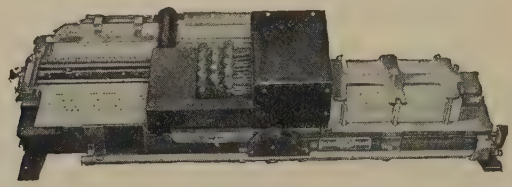


FIGURE 120.—Electric Key Punch with Duplicating Attachment

punching a master card with these common data and inserting it in the space provided for it at the upper left hand corner of the machine as shown in the cut.

By means of electrical contact the holes punched in this master card are automatically reproduced in all the other cards that run through the machine, these cards being fed automatically from the card magazine to the punching device. If desired, however, cards of different form or color may be fed directly into the punch without removing or disturbing those in the magazine. When individ-

ual data are to be recorded the machine stops and these data are then punched by the operator. With no master card in position the machine is used as a regular electric key punch.

The electric key punch with duplicating attachment is designed to duplicate with absolute accuracy any punched card in its entirety, or the punching of any portion or portions of the card. In using codes, for example, such as those for Branch, Customer, State, County, etc., the clerical labor of inserting codes on the original records prior to punching is eliminated by this machine through the use of pre-punched code cards. The data recorded on the code card are automatically duplicated on each of the individual record cards (to which the code applies) at the time they are key-punched. The necessity for verifying the duplicated cards is avoided.

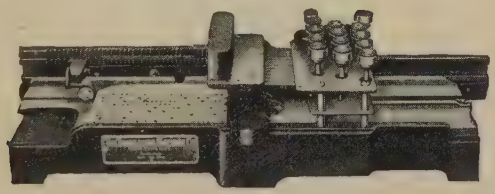


FIGURE 121.—Verifying Key Punch

The Verifying Machine.—By means of a machine generally similar to the key punch machine, the cards are verified. It is operated by means of keys in the same way as the key punch and is called the verifying machine (Figure 121 above). Cards which are punched by the gang punch can be verified by holding a bunch of them to the light, or running a small rod such as a knitting needle through the holes.

The Sorting Machine.—When the data have been placed on the cards by means of the punched holes, the record is complete. The remaining machines of the system are employed in using the record. The sorter, by grouping the cards, classifies the data and the tabulator summarizes and tabulates them.

The *sorter* (Figure 122 below) is much more complex in construction than the punching machines but is simple to operate. By means of this electrical machine which operates automatically the cards are separated and thrown into separate pockets. In this

way the cards may be sorted into groups to furnish any classification of the facts which have been recorded. In the analysis of

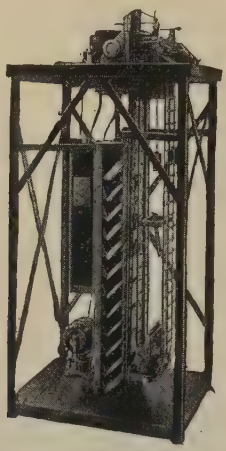


FIGURE 122.—Electric Sorting Machine

sales data, for example, the cards may be sorted by salesmen; each of these groups may then be sorted by classes of commodity; these again by customers; and so forth. A tremendous amount of

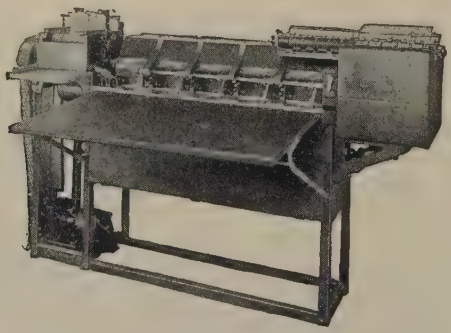


FIGURE 123.—Electric Accounting Machine

classification can be accomplished by this process in a relatively short period of time.

The Tabulating Machine.—Primarily, the tabulating machine (Figure 123 above) called the Electric Accounting Machine,

summarizes the data which have been classified by the sorter. Here the cards are simply fed into the machine and the totals are registered and recorded automatically.

This machine makes printed records and is so constructed as to permit the printing of the data on standard office records and reports, thus obviating the necessity of typing the data that have been analyzed.

Service From the Hollerith System.—The description of the method of operation given above is of course not intended to direct the reader how to operate the machines. Such directions will be freely given by the Tabulating Machine Company on request either by correspondence or by interview and they will be accompanied by demonstrations of the machines if convenient.

It will be seen from the preceding discussion that the chief advantage of the Hollerith system lies in the tremendous volume of work that can be done in a relatively short period of time. There are so many variable factors, such as the character of the data to be recorded, the condition of the records from which the facts are taken, and the skill of the operators, that it is difficult to make a definite statement regarding the volume of work. In fact, so varying are the results that the Tabulating Company is changing its policy relative to the publication of definite figures, although it has endeavored in the past to determine figures representing average production for the various machines in the hands of operators of average skill. As these companies offer service throughout the country it is always preferable to consider with them the specific problems that we have to solve. The great economy comes from the sorting and tabulating and these are the only processes involved after the records have once been made.

It is clear that for any great volume of work, accounting or statistical, the tabulating system should be given most careful consideration. The machines are rather expensive, however, and care must be exercised in introducing the system lest the volume of work be not great enough to keep the machines fairly busy. Part-time production naturally increases the relative cost and may make it unwise to adopt the system. If the volume of work will justify it there is no question as to the advisability of using this type of

mechanical device. The general service of these machines is aptly expressed by an executive of a large industrial organization in Chicago who had been using them—"Many investigations of special problems heretofore unattempted because of irksome detail and expense, can be carried on with care and exactness."

Frequently, statistical and accounting work can be combined to furnish enough work to warrant the use of the tabulating machines or in small concerns several different companies may cooperate in maintaining a single office using codes to preserve the confidential character of the data. From the practical point of view, it is always advisable when a considerable volume of statistical work is to be done, to take the problem up with representatives of the tabulating machine companies and consider the possibilities of this system. The system is especially serviceable where recorded data are to be used frequently. In such cases it not only analyzes the data economically but it produces the results much more quickly than any other method. For speed and accuracy this general method of recording, analyzing, and tabulating statistical data is not surpassed. The principal question is in relative cost and this can be decided only on the basis of the facts of the particular situation.

The Slide Rule.—The slide rule, or what is sometimes called the "slip-stick," is a very useful mechanical device of which every statistician should have a practical knowledge. It should be emphasized that, contrary to popular opinion, the slide rule is simply constructed and its practical operation can be readily understood. The principles on which it is based and the rules for its practical operation are simple. The theory of the slide rule is developed from the elementary principles of logarithms and a complete understanding of its nature presupposes familiarity with these elementary principles. However, we may become proficient in the operation of the slide rule without knowing the fundamental principles of its construction, just as we can drive an automobile without knowing all of the principles involved in its construction. In this chapter we shall outline briefly the nature of logarithms and how the slide rule is developed therefrom; and we shall also

indicate the simple rules for the practical operation of the slide rule.

The Theory of Logarithms.—The word "logarithm" means etymologically a ratio or relationship between numbers. Hence, in building a system of logarithms we express a certain relationship between one number which we take as the base of the system and all the numbers severally whose logarithms we assume to develop. This relationship involves the expression of each of the numbers whose logarithms we undertake to determine as a power of the number taken as the base.

Hence, the logarithms of a number may be defined as the exponent of that power of the base which is equal to the number. Thus, if 10 be taken as the base, since 10^2, 10^3, and 10^4 equal respectively 100, 1,000, and 10,000, 2 is the logarithm of 100, 3 of 1,000, and 4 of 10,000, or, in other words, 100 is the second power of 10, the base, 1,000 the third, and 10,000 the fourth. Any positive number may be taken as the base of a system of logarithms but as a matter of practice what is known as the Common System, of which 10 is the base, is so general that we use the logarithms of this system without indicating the base. It is only when some other base is used that we need specify the base in practical work. In this discussion we shall recognize only the Common System.

Composition of Logarithms.—If a number is not an exact power of 10 its logarithm can only be expressed approximately and consequently we have logarithms containing decimals, such as 0.30103 the logarithm of 2; 0.95424 the logarithm of 9; and 1.32222 the logarithm of 21. Hence, we recognize two separate and distinct parts in the composition of the logarithm:

1. *The Characteristic*—This is the integral part of the logarithm—*i.e.*, the part to the left of the decimal point—and may be either positive or negative. In the logarithm of 21, for example, 1 is the characteristic.
2. *The Mantissa*—This is the decimal part of the logarithm and is always kept positive in order that the mantissa of the logarithm of all numbers expressed by the same digits in the

same order may always be the same. Thus, in the logarithm 21, .32222 is the mantissa.

How to Determine the Logarithms of a Number.—To facilitate the determination of logarithms, "Tables of Logarithms" have been published by which the logarithms of any given number or the number corresponding to any given logarithm may be found. The mantissa of the number will be determined directly from the table by following the accompanying directions. The characteristic will be determined by simple rules, which are fully explained in the "Tables" just referred to and which are given in connection with the logarithms in Appendix D.

Properties of Logarithms.—Logarithms are very serviceable in certain arithmetic computations, particularly where large numbers or large groups of numbers are involved. The following outlines the method of performing the computations:

1. *Multiplication*—The sum of the logarithms of two or more numbers is the logarithm of the product of the numbers. Hence, to multiply by means of logarithms we (1) find the logarithms of the numbers to be multiplied together, (2) summate these logarithms and (3) find the number corresponding to the logarithm resulting from No. 2. This is the product desired. Thus log 2(0.30103) + log 3(0.47712)= log 0.77815, which is the logarithm of 6 as determined from the table of logarithms.
2. *Division*—The logarithm of a quotient is equal to the logarithm of the dividend minus the logarithm of the divisor. Hence, to divide by means of logarithms we (1) find the logarithms of the dividend and the divisor, (2) subtract the logarithm of the divisor from the logarithm of the dividend, and (3) find the number corresponding to the logarithm resulting from No. 2. This number is the desired quotient. This can be illustrated simply by reversing the computation given in No. 1 above.
3. *Powers*—The logarithm of any power of a number is equal to the logarithm of the number multiplied by the exponent of the power. Hence, specifically, we (1) find the logarithm of the number, (2) multiply this logarithm by the exponent of the power to which the number is to be raised, and (3) find

the number corresponding to the logarithm resulting from No. 2. This is the desired power of the number. Thus log 3(0.47712) multiplied by 3 gives log 1.43136 which is the logarithm of 27, the cube of 3.

4. *Roots*—The logarithm of any root of a number is equal to the logarithm of the number divided by the index of the root. Specifically, we (1) find the logarithm of the number, (2) divide it by the index of the root, and (3) find the number corresponding to the logarithm resulting from No. 2. Reversing the operations of No. 3 above will illustrate the process.

Further discussion of logarithms will be found in the references listed on page 514. In this connection it should be noted that logarithms themselves are very serviceable in cases where extensive computations of large numbers are to be made.

Construction and Operation of the Slide Rule.—The slide rule is a simple mechanical device designed primarily to facilitate the computations by means of logarithms outlined above. It consists essentially of graduated scales which are constructed to represent the logarithms of numbers by means of linear distances.

Bearing in mind that we want to add and subtract logarithms, thus accomplishing the processes of multiplication and division, the simplicity of the fundamental mechanical principles will be apparent. If, for example, we have two linear distances ("a" and "b" in Figure 124 below) which respectively represent two numbers according to a common scale, then the combined length of these two distances ("c" in Figure 124) will be the linear distance representing the *sum* of the two numbers to the same scale and the difference between these two distances ("d" in Figure 124) will be the linear distance representing the *difference* between the numbers to the same scale. This is the simple mechanical principle underlying the slide rule and the mechanism is designed so as to facilitate the *combining* of linear distances and of taking their *differences.*

This is accomplished by constructing two identical scales to a common base and placing them side by side in such way that they will readily *slide* past each other, thus permitting any two graduations on the separate scales to be brought together. Thus, in

"A" of Figure 125 below, the graduations on scale "c" lie directly opposite corresponding graduations on scale "d"; in "B" of the

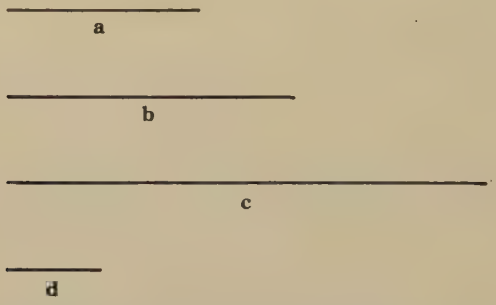


FIGURE 124

If "a" represents 2, then by the same scale "b" would represent 3, "c" (which is constructed by joining "a" and "b") would represent 5, and "d" (which is the difference between "a" and "b") would represent 1.

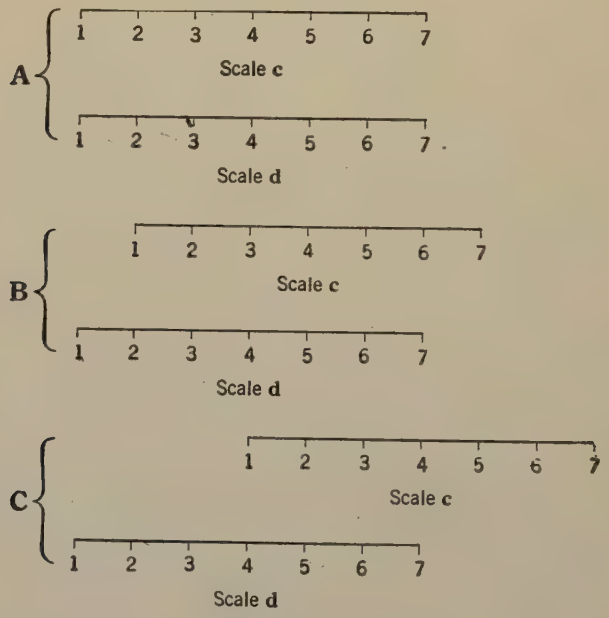


FIGURE 125

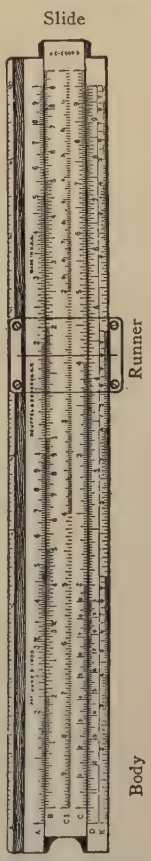
same figure, scale "c" has been pushed to the right and the graduations no longer correspond, 1 on "c" being opposite 2 on "d,"

2 on "c," being opposite 3 on "d," and so forth; and in "C" of the same figure, scale "c" has been pushed even farther and 1 on "c" now corresponds to 4 on "d" with corresponding adjustments in other designations.

It is clear, therefore, that by moving the slide we can bring any point on scale "c" opposite any point on scale "d" and this enables us to add or subtract any distances represented on the scales. If, for example, we want to add the distance between 1 and 2 on scale "c" and that between 1 and 3 on scale "d," we simply move the slide until 1 at the left end of scale "c" is precisely opposite 3 on scale "d" and the distance between 1 at the left end of scale "d" and a point on scale "d" precisely opposite 2 on scale "c" is obviously the desired distance—*i.e.*, the sum of distances 1 to 3 on "d" and 1 to 2 on "c."

Likewise, if we wish to get the difference between these two distances, we place 2 on "c" at 3 on "d" and the distance between 1 at the left end of "d" and a point on "d" precisely opposite 1 on the left end of "c," represents the difference.

In order to facilitate computations by this simple process, the slide rule carries three sets of corresponding scales, "C" and "D," "A" and "B," and "A′" and "B′," as shown in Figure 126. It consists of three separate parts, as indicated in Figure 126, the *body*, which is the fixed part, the *slide*, which is the movable or sliding part that fits into and slides in the groove in the body, and the *runner*, which consists of a glass set in a frame and carrying a hair-line on its under side. The actual operation of the rule is described on page 511.



(Courtesy Keuffel & Esser Co.)

FIG. 126

How to Read the Slide-Rule Scales.—Perhaps the greatest difficulty to the beginner in the use of the slide rule arises in connection with the reading of the scales. The facts (1) that the distances between consecutive numbers on the scale are unequal, and (2) that the points on the scale may represent different values.

tend to create confusion in his mind. Regardless of these facts, however, the scale is read in essentially the same manner as scales which do not have these complications. For convenience, the ends of the scales are called indexes, the 1 at the left end being known as the left index and the 1 at the right end, the right index as shown in Figure 126 above.

The essential nature of the graduated scale is developed in Chapter XXVIII. It will be noted there that the scale is built up from a basic unit length which represents a certain numerical value. It will be noted also from the discussion referred to that we interpolate in this simple scale by taking such part of the total distance between any two main designations as the excess of the number to be recorded over the first scale designation is of the difference between the two main designations or, *vice versa,* if we wish to determine the value of any given point between two main designations on the scale we take such part of the scale unit as the distance from the first scale designation is of the total distance between the two. Thus, in Figure 125 above, 2.5 is located on the scale just half-way between 2 and 3, 3.33 1/3 one-third of the distance from 3 to 4 or, *vice versa,* the point half-way between 2 and 3 represents 2.5 in value and that lying one-third of the way from 3 to 4 represents 3.33 1/3. In strict theory, interpolation on the logarithmic scale is somewhat different from that on the arithmetic scale as noted on page 333.

Now, this fundamental method of locating values on a scale is followed in reading the slide-rule scales but the nature of the scale tends to conceal the principle and to make the reading of these particular scales more difficult.

Unequal Distances.—The scale of the slide rule is developed on the basis of the Common System of logarithms whose base is 10 and the irregularity of the graduations is explained by the fact that the scale is graduated to represent the logarithms of the numbers from 1 to 10 as shown in Figure 78, page 328. To facilitate the use of the scale, the numbers which the logarithms respectively represent are substituted for the logarithms as shown in Figure 79, page 328. Thus, the point designated by 2 on the scale, for example, does not in reality represent 2 but the logarithm of 2,

This affords more convenient scale units and obviates the necessity of finding the logarithms corresponding to numbers and *vice versa.* In this connection, the reason why the scale begins with 1 instead of 0 will be apparent—*i.e.,* 1 really represents the logarithm of 1 which is 0 and so the true scale values are the logarithmic values of 0 to 1.

From the foregoing, it will be seen that in reality there is no irregularity in the values of a given distance. While in Figure 79, page 328, apparently 1, the difference between 1 and 2, is represented by the distance from 1 to 2, while the difference between 3 and 4, also 1, is represented by the much smaller distance from 3 to 4, the fact is that the distance between 1 and 2 does not represent the difference between 1 and 2, but the difference between their logarithms—*i.e.,* .301 minus 0—and likewise the distance from 3 to 4 represents the difference between the logarithms of these numbers—*i.e.,* .602 minus .477 or .125. In other words, the distance from 1 to 2 represents roughly .301, and that from 3 to 4, roughly .125, which is in accordance with the fundamental principle of plain scale construction that equal distances must represent equal values at any place on the scale.

The foregoing is merely for the purpose of showing that the graduations were made strictly in accordance with the principles of scale construction with which we are familiar. It will be observed, however, that when we use the substituted numbers instead of the logarithms *equal distances represent equal ratios* instead of equal differences, the distance between 2 and 4, for example, being equal to the distance between 4 and 8, or, the ratio between 3 and the value at a point a half-inch to the right of 3 is the same as the ratio between 5 and the value at a point a half-inch to the right of 5. In other words, we have a logarithmic scale such as was discussed in Chapter XXXI.

Different Values.—The second difficulty noted on page 507, that any given point on the scale may represent different values, is also due to the fact that the scale is logarithmic in character. It is true of the logarithmic or ratio scale that the integrity of the scale is maintained so long as the values assigned to the several

graduation points bear the same ratio to one another. That is, any values, for example, may be assigned to the points marked 1 and 2 which bear the same ratio to each other as 1 and 2 bear to each other, it being necessary, of course, to change the designations on the remaining portion of the scale proportionately. Hence, it is clear, that the series of numbers from 1 to 10 which are listed on the scale may be multiplied or divided by any number without affecting their relative values or impairing the character of the scale. This means, concretely that we may multiply or divide the scale designations, item by item, say by 5 or 10. In short, we may multiply or divide by any number so long as we give the same treatment to each of the designations on the scale.

So the apparent difficulty becomes indeed a great convenience. It enables use conveniently to read from a single scale, numbers which differ greatly in value, the point 25 on the scale, for example, being read .25, 2.5, 25, 250, 2,500, 25,000, etc., depending on the divisor or multiplier which we use to change the designations. Now, for convenience, in practice, we usually keep the scale designations in multiples of 10. Thus, the scale may read 10, 20, 30, etc.; 100, 200, 300, etc.; 1,000, 2,000, 3,000, etc.; or, going in the other direction, .1, .2, .3, etc.; .01, .02, .03, etc.; .001, .002, .003, etc.; depending on the size of the number we wish to read. Confining the scale adjustments to multiples of 10 enables us readily to make adjustments mentally as we perform the operation. In reading an integral number we simply add as many ciphers to each main scale designation as there are digits, less 1, in the number to be read: if a decimal number we give the left index the same decimal designation as the number to be read. For example, if we wish to read 75 and 755, the designation at the point 7 will represent 70 and 700 respectively in the two cases. It will be kept in mind in this connection that the right index is always 10 times the left index.

It should be noted, furthermore, that the decimal point is disregarded in locating a number on the scale,—125, 12.5, 1.25, and .125, for example, being located at exactly the same point. The location of the decimal point is taken care of in another way as explained on page 512 below.

Practical Operation of the Slide Rule.—As stated above, a knowledge of the theory of the slide rule is not essential to proficiency in its use. We can simply take certain things for granted and be guided by a few simple rules noted below, detailed discussion of which will be found in the instructions accompanying the slide rule.

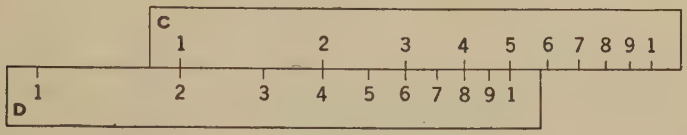


FIGURE 127

To multiply 2 by 3 we place 1 on "C" at 2 on "D," and at 3 on "C" we read the product on "D"—*i.e.*, we take the sum of the distances 1 to 2 on "D" and 1 to 3 on "C" which as the scales are set is 1 to 6 on "D."

To divide 6 by 3 we place the divisor 3 on "C" at the dividend 6 on "D," and at 1 on "C" we read the quotient 2 on "D."

It will be recalled in this connection that multiplication involves the addition of logarithms and that division involves subtraction. The operation of the slide rule will be illustrated by using scales "C" and "D."

1. *Multiplication*—To multiply, we place the *proper* index of "C" directly above one of the numbers on "D," and directly under the other number on "C" we read the product on "D." We use whichever index will let the number on "C" fall above the scale "D" and not beyond it. If we use scales "A" and "B" we can use the one index all the time and read the results on the other set of scales. Figure 127 above illustrates the multiplication process. The method of determining the decimal point is discussed below. The use of a constant multiplier or the process of continued multiplication will be apparent; they are discussed in books on slide-rule instruction.
2. *Division*—Here we wish to subtract the distance representing the divisor from that representing the dividend. This is accomplished by locating the divisor on "C" and the dividend on "D," setting the one directly over the other, and directly under the proper index on "C" finding the quotient on "D." The meaning of "proper" is indicated under "Multiplication" above and the method of determining the decimal point is developed below. The process of division is illustrated by Figure 127 above.

3. *Powers and Roots*—The squares of numbers on scales "C" and "D" are located directly above on scales "A" and "B" since the latter scales are only half as long as the former. Reversing the process will enable us to determine the square root. Further directions as to powers and roots will be found in treatises devoted exclusively to the subject.

This discussion will give the reader a general idea of the practical operation of the slide rule. For finer points in the operation of the slide rule as well as for more complete description of its construction the student is referred to the instructions which generally accompany the rule and to references on page 514.

Locating the Decimal Point.—It is obvious that the slide rule does not indicate the decimal point in either multiplication or division. This must be determined independently from the slide-rule operation. The following rules are sometimes given for locating the decimal point:

Multiplication—The number of digits in the product is equal to the sum of the number of digits in the two factors, except where the slide projects on the right, in which case one must be subtracted from the sum of the number of digits. This rule will apply either to the "C" and "D" scales or the "A" and "B" scales. It should be noted that the rule will not work in problems in which the reading is the right index of the slide such as 25×4 or 125×8, and so forth. This is an exception to the rule.

Division—The number of digits in the quotient is equal to the number of digits in the dividend less the number of digits in the divisor, except where the slide projects on the right, in which case one must be added to the difference between the number of digits. This rule will apply to both the "C" and "D" and the "A" and "B" sets of scales.

Use of the Slide Rule.—These brief suggestions are merely for the purpose of indicating the broad methods of procedure in the operation of the slide rule. For special points of operation such as the method of doing continued multiplication or division or of handling a constant multiplier, dividend, or divisor, the

student is referred to the instructions which accompany the slide rule. A description of the additional graduations on the rule and of the method of its construction are also given in those instructions and in the references listed below.

The results obtained from the ordinary slide rule are not of sufficient degree of exactness to permit its general use in statistical computations. With the full-length scales of the common ten-inch slide rule, accurate readings can be readily made to three places while with the longer rules which are graduated more finely, readings can be made to four places. In any case, experience in the use of the rule enables one to interpolate more accurately so that figures of four or five places may be read with reasonable accuracy. However, because of this limitation on the exactness of the readings, the use of the slide rule in statistical work is confined for the most part to rough computations of percentages and ratios, particularly in checking important computations in final statistical statements before they are released.

There are specially constructed instruments of this general type such as Webb's Stadia Slide Rule and Thacher's Calculating Instrument which yield exact readings to a greater number of places. These are equivalent to straight rules several feet in length and consequently furnish much finer graduations.[1]

Summary and Conclusion.—In addition to the foregoing there are certain miscellaneous aids more general in character, such as patented card record systems, mimeograph, multigraph, photostat, blue-print, offset process, lithoprint, and addressograph machines, tables of squares and cubes, and so forth. These devices, however, are of general service to the business organization and consequently do not call for special treatment here. In view of the fact, however, that we usually require several copies of statistical statements, it is advisable to make careful examination of the different methods of reproduction.

The foregoing brief account of aids in statistical work is designed primarily to suggest the possibilities along this line. The

[1] The Student's Slide Rule put out by Keuffel and Esser Co. (price $.75 and $1), is particularly serviceable to teachers. It is cheaply constructed but is entirely satisfactory for class-room work. For class-room work also, a so-called *demonstration rule* may be obtained from the same company. This is a large rule, 6 or 8 feet long, with prominent markings and is especially serviceable for instruction purposes.

volume of work required in the modern business organization necessitates the use of these mechanical devices. The statistician should always keep closely in touch with the development of improved mechanical aids and select those which may be serviceable in the particular case. Generally speaking, it will be advisable to centralize the work of the organization at least to the point where advantage can be taken of the economies that can be realized through the use of these devices.

It should be noted in this connection that the agencies selling these devices will generally cooperate freely in working out the use of the machines in any special case. Representatives will call in person, if convenient, or assistance will be offered through correspondence.

It will be advantageous for the teacher to equip his statistical laboratory with a number of these machines if funds are available. This will enable him to increase the demands on his students for practical work and will also acquaint the students generally with the operations of these devices and the services which may be derived from them. If funds are not available for such purposes, it is sometimes possible to have the machines placed in the class room for a short period of time for demonstration purposes, in case distributing points for the machines are located in the community.

References

American Digest of Business Machines (latest issue). American Exchange Service, Chicago.

Barlow, Peter. Barlow's Tables of Squares, Cubes, Square Roots, Cube Roots, and Reciprocals of all Integer Numbers up to 10,000 (1914).

Blaine, R. G. The Slide Rule, 5th Ed., 1921.

Bowser, E. A. Logarithmic Tables, 1909.

Hinckley, T. C., and W. W. Ramsay. The Slide Rule, 1910.

King, Willford I. Elements of Statistical Method, 1912, Appendix A.

McGraw-Hill. Six Place Tables, 1922.

Rugg, H. O. Statistical Methods Applied to Education, 1917, pp. 66-72.

U. S. Bureau of Labor Statistics. Bulletin 326—Methods of Procuring and Computing Statistical Information.

Wells, Webster. Six Place Logarithmic Tables, 1907.

CHAPTER XLVI

STATISTICS AND ACCOUNTING

Statistics and Accounting Distinguished.—While the actual delineation of the work of a statistical organization in a practical business concern involves the consideration of the line of demarcation between the statistical organization and the operating departments of the company, it is of chief concern in considering the relation between statistics and accounting. This is due to the great similarity between these two functions, each of them dealing more or less exclusively with the assembling, recording, analyzing, and interpreting of numerical facts.

A great variety of ideas have been offered in distinguishing the two, ranging from the conception that accounting is but a specialized form of statistics to the opposite extreme that accounting includes the assembling, recording, analysis, and interpretation of all types of business facts. A review of these conceptions in detail would contribute little to the purpose of this discussion. It is sufficient to say that generally distinctions are recognized in (1) the purpose, (2) the type of data, and (3) the methods.

Purpose.—In the development of scientific business management, the field of accounting has been greatly broadened. In former times, the aim of accounting was confined almost exclusively to the development of the *balance sheet* and *income statement* but more recently, as an aid in the management of the business, attention has been paid to the analysis and interpretation of the detailed facts from which these statements are prepared. Refined classifications are maintained and averages, ratios, and per unit figures are developed to reflect the bearing of these financial facts on the practical problems of management. The recent books by Bliss on "Operating Ratios" [1] and "Management Through Accounts," [2] indicate these new developments.

[1] Financial and Operating Ratios in Management, J. H. Bliss. Ronald Press Co., N. Y.
[2] Management Through Accounts, J. H. Bliss. Ronald Press Co., N. Y.

In spite of these refinements, however, the principal function of accounting still is to record the transactions of the business and to develop records and prescribe methods which enable us to reflect the financial status of the business at a given date and its profit or loss over a given period of time. This is well stated by M. B. Ignatius in an article on the "Relation of Statistics and Accounts," in *Industrial Management* for October 1918:

"In spite of many refinements, the primary purposes of accounting still relate to the treasury—the earmarking of sources and destination in order that funds and claims may be properly identified, collected, and made to discharge the liabilities of the business. Its chief aim is still that of resolving all transactions into net income or profit and loss."

This is distinctly an accounting job just as making purchases, manufacturing products, and selling goods are functions respectively of the purchasing, producing, and marketing departments. It is the essential purpose of accounting and no one would consider it a statistical job any more than he would consider making purchases, manufacturing products, or selling goods, statistical jobs. In this way, then, the purpose of accounting differs from that of statistics.

Type of Data.—It is obvious, therefore, that accounting is confined essentially to the field of financial data—that the accounting unit is almost exclusively dollars and cents. On the other hand, statistics deals with any and all types of data—personnel and operations as well as financial—which may be expressed numerically. This more comprehensive conception of statistics is also apparent from the fact that statistics are used extensively in other fields—education, sociology, anthropology, biology, etc.—while accounting is confined for the most part to those processes which involve dollars and cents.

Furthermore, in practical work in a business concern, accounting is confined generally to internal facts while statistical developments will involve a great body of facts from the external field as indicated in Chapters IV, VI, and VIII.

Methods.—A distinction is also recognized between accounting and statistical methods. Each is governed by a body of special

principles designed to serve a specific end. The accounting principles are not covered in a treatise on statistics nor the statistical principles included in accounting treatises.

The principles and practices of collecting statistical data discussed in Chapters XII to XX are generally not considered important in accounting since accounting data are taken from formal accounting records and difficult problems of collection do not arise. Likewise, analytical methods in accounting are confined primarily to classification and simple computations, the methods of sampling, frequency classification, special types of averages, dispersion, skewness, and correlation—all of them effective statistical devices—being almost unknown in accounting.

Furthermore, in accounting we are interested primarily in the values of the figures in dollars and cents and we, therefore, strive for accuracy to a penny. In statistics we are interested more in the conclusions which the figures indicate and consequently can make use of approximations and less minutely accurate methods generally.

A Practical Conception.—Having noted the rough distinctions given above, let us proceed immediately to forget them. In practical work, fine distinctions between the two are unnecessary. The important point is that we want the facts collected, analyzed, and interpreted and the two bodies of methods will be used wherever required to accomplish this end: what we call the organization that uses these methods is of minor concern. In this respect, we may well adopt the burden of the oft-quoted definition of psychology, that "psychology is the science which treats of the problems that the psychologist studies." In this phase of business management we are not nearly so much interested in what we call the work as we are in having the full force of the facts, properly analyzed and interpreted, brought to bear on the problems of management.

The "Fact" Department.—As a practical procedure, therefore, we should recognize the similarity between these two functions in forming the organization to carry them on: we should have a general department to include both statistical and accounting work. This "fact" department may be called by whatever

name may seem appropriate in the particular company—accounting, statistical, comptroller's, statistics and accounting or what not.

The Accounting and Statistical Organizations.—The detailed organization of this "fact" department may take a number of different forms each of which may be preferable under particular conditions. As a concrete method of procedure, however, whatever may be the nature of the organization, there should be a unit or section whose chief function is the development of the balance sheet and income statement with all the supplemental work that this implies. It will involve full and complete records covering revenues and expenditures, and the devising of methods to satisfy accounting requirements regarding their execution and delivery to the accounting organization. This work is vitally essential to the business and nothing should be permitted to interfere with it.

However, this is as far as the distinction need be made. All other work in the "fact" department should be done without regard for its characterization as between statistics and accounting and likewise without characterization of the group of employes who do it.

For practical purposes, we may call the organization, which is responsible for the strict accounting work noted above, the accounting organization; we may call the organization which is responsible for the other work in the "fact" department, the statistical organization. This distinction, however, is purely for convenience. It is not designed to indicate definitely what work shall be done exclusively by one organization and what by the other. It is merely the same basis on which we distinguish the statistical organization from the operating departments of the company. The sales department, for example, is the sales department because it is responsible for selling the company's products but sales statistics may be handled by the sales department or by the statistical organization. Likewise, the distinguishing function of the accounting department is the development of the balance sheet and income statement but here also the accounting organization may do certain statistical work just as may the sales or any other operating department.

Cooperation between the Statistical and Accounting Organizations.—As noted at the outset of this chapter, we are primarily interested in the cooperation between the statistical and accounting organizations rather than their distinction. They must work together in a number of ways if the best service is to be obtained. The regular accounting records should be used to the fullest extent in collecting the raw data required in the entire "fact" department, whether or not these data are used specifically in the development of the formal accounting statements. The use of these records in collecting internal facts is discussed in Chapter X. Although nothing should be permitted to hamper seriously the essential service of these records, nevertheless additional data of other types and for other purposes can generally be recorded on them without detracting from their services as strict accounting records. However, operating records required for statistical work but not for strict accounting should come directly to the statistical organization.

There is certain other work which the accounting force will be better equipped to do, and wherever this is the case such work should be done by them. This will apply particularly to analytical reports which are derived from the data required in the preparation of the formal statements, since all figures of this type should check with the formal accounting records. The amount of additional work to be done by the accounting organization will depend on the facts of the given case. In any event, it will be confined for the most part to financial facts but will generally be pretty extensive within this field.

Obviously, since the statistical and accounting work is very closely related, the two forces should be generally familiar with both types of work. While the accountants and statisticians will be specially qualified each for his own type of work, the best service will be obtained when both are generally familiar in the entire field of the "fact" department. It is particularly important that the accounting force be familiar with statistical methods and requirements when the accounting records are being devised in order that the statistical service from these records may be fully appreciated and statistical requirements provided for. It is a fair

assumption that the accountant of the future who is not pretty well familiar with statistical requirements will find little demand for his services, particularly in installing accounting systems.

This type of organization will also permit the most efficient and economic use of mechanical aids. As the statistical and accounting work is done in such close cooperation, a common service organization can be used to do the computing, drafting, and clerical work for both the accounting and statistical organizations.

References

Cole, W. M. Accounts, Their Construction and Interpretation, (Revised Ed. 1915), Ch. XI.

Franklin, B. A. Cost Reports for Executives, 1913, Ch. VII.

Harrison, G. C. Cost Accounting to Aid Production, 1921.

CHAPTER XLVII

ORGANIZING FOR A PROGRAM OF STATISTICS

Place of the Statistical Organization in the Company.—At this point we shall consider briefly the building up of the statistical organization—its place in the company, its personnel, its equipment, and its cost.

The statistical organization may be located at various places in the company, according to the way it has grown up, and good statistical service is being rendered in practical business concerns by statistical organizations having different places in the company. Generally, however, it is preferable to have the statistical organization located in the accounting department with the chief statistician reporting to the head of that department. This is because of the fact that the accounting department in its routine functions maintains records of a great amount of statistical detail and as a matter of fact the system of basic statistical records will perhaps be preferably built up through an expansion and elaboration of the regular accounting records. Furthermore, as noted in Chapter XLVI, in much of the statistical work the data must check with the official figures of the accounting department and the close association of the two organizations facilitates cooperation along this line. An advantage may, likewise, be realized through cooperation in the use of mechanical aids, particularly the Hollerith system.

It is sometimes contended that the statistical organization should be associated with the office of the president or operating vice-president, since the service to be rendered is company-wide in scope and should not be narrowed by being identified with a single department. It is further contended that since the statistical organization must obtain the detailed data from the several operating departments, the authority which the office of the president or operating vice-president lends to the organization is essential to proper service. As a matter of fact, however, while the

authority of these officers is of course necessary for the taking of any important steps in the development of the program of statistics, it can be obtained quite satisfactorily with the statistical organization located in the accounting department.

Centralized versus Decentralized Statistical Organization. The statistical organization may be centralized or decentralized in form. In the one case, (1) the chief statistician is immediately and directly in charge of the complete program of statistics, (2) the entire statistical force is centralized under his direct supervision, and (3) a comprehensive system of primary basic data records is maintained in the statistical organization. (See Figure 128 below.)

In the other, (1) the chief statistician only remotely and indirectly supervises the statistical work, (2) only a supervising nucleus of the statistical force is located in his office, the bulk of the force being distributed among the several departments of the company in the form of operating units who are under the direct supervision of their respective departments, and (3) primary basic data records are to be found for the most part only in the several departments of the company with or without special organization or systematization to facilitate statistical work. (See Figure 129 below.)

Which of these forms of organization is preferable will depend to some extent on the organizational character of the particular establishment that the program of statistics is designed to serve. Frequently, the policy of the business is one of decentralization and this may at times be a controlling factor in determining whether the statistical work shall be centralized or decentralized. There are, however, certain fundamental considerations which play an important part in deciding the type of organization.

The business concern may be a simple organization as in the case of most smaller companies, or it may be made up of a number of associated companies as in the case of some of the larger American corporations. Furthermore, all the operations and employees of the concern may be grouped together under one roof or in one locality, or through branches or geographical divisions they may be scattered over a wide territory with the several parts of the

organization widely separated over the country or even over the world. Where the business organization is widely separated it is of course imperative that the statistical work be likewise decentralized. The question of the comparative value of the centralized and decentralized type of statistical organization has to do specifically with the simpler type of organization where the operations and forces are concentrated at least in one locality.

The Centralized Organization.—The following points are generally noted to indicate that the centralized organization is preferable:

1. *Accuracy*—The centralized statistical organization is better prepared to guarantee accurate results since it keeps under the immediate supervision of the head of the statistical organization the entire work of developing statistical statements, from the primary basic records to the completed study. Better facilities are provided for checking individual operations and errors can be detected and corrected with less expenditure of time and effort. The larger and more experienced corps of personnel, the closer co-operation in each phase of the work, and the more efficient supervision, tend to produce more accurate results than are obtained in the decentralized organization.
2. *Relieves the Departments*—The centralized statistical organization conforms to, and the decentralized statistical organization conflicts with the principle of management which prescribes that the operating department of a business organization should be relieved of all burdens of the business except those associated directly with operations. Disregard for this principle is disapproved not only by the department heads but also by the general executives of the company. An operating department is charged with the responsibility of accomplishing certain specific things and the general policy of the management usually insists that burdens which fall outside its definite field of operations be reduced to a minimum. Accordingly, in the statistical work of the business, the operating departments should be relieved of any material burden not only in the development of statistics for the use of the general officers of the company but also even in the development of statistics for the use of the departments themselves. In a decentralized

statistical organization, where the burden of statistical operations falls upon the operating departments of the company, the chief statistician is constantly in hot water lest his demands upon the operating departments be too great. The probable burden upon them is an important factor in deciding whether or not a specific statistical study should be undertaken.

The centralized statistical organization, on the other hand, relieves the several operating departments of the company of practically all the burden involved in the statistical work of the company. Comprehensive primary basic records are maintained in the centralized organization and all forms of statistical work—secondary basic records, routine reports, special studies, and so forth—can be undertaken without disturbing the operating departments. The only burden on them is that involved in providing the statistical organization currently with copies of such primary records of the business as may be required. This burden is comparatively negligible.

3. *Statistics a Main Job*—When the statistical work is centralized the entire development is under the immediate supervision of the chief statistician and his staff, whose chief function in the company is to do statistical work. When this statistical work is done in the several operating departments—when it is decentralized—the supervision is done by those whose main jobs are not statistical compilations but operating work. The immediate supervision of the chief statistician is far preferable. This is his main interest. He studies statistics and equips himself to do the best kind of work in this field. The difference between having the work supervised by the chief statistician's office, on the one hand, and by the operating departments of the company, on the other, is just the difference between having the statistical work a main job and a side issue.

4. *Uniformity*—A centralized statistical department insures uniformity of treatment both in the organization of the work and in its detailed development. Any one who has undertaken to have statistical data prepared along identical lines by different persons—individuals or groups—knows how unsatisfactory the results may be even though the supervision of the work has been most careful. Inevitably, questions of detail arise in the preparation of the material and, if the final

developments are to be comparable, these detailed questions must be uniformly disposed of. In the centralized department such decisions are made by a single authority rather than by several as in the case of the decentralized department.

Furthermore, in the decentralized statistical organization the systematic organization of basic data which is the first step in the centralized department and which makes possible a complete and comprehensive statistical service, is lacking. The absence of this comprehensive foundational material greatly handicaps the chief statistician and definitely limits the service which he can render to the company. In the decentralized department, statistical studies are limited to the data which are maintained uniformly by the several departments. This fact frequently eliminates very important studies. With the comprehensive system of basic data records built up in the centralized department, any statistical study which may be required can be developed. Centralization means uniformity and comparability in the material; these make possible combinations and summarizations in almost any form.

5. *Avoids Duplication*—The centralized statistical organization tends to avoid duplication both of effort and working force. In the decentralized type of statistical department, a miniature statistical organization exists in each operating department of the company and many operations are performed severally by these units which might be handled as a single operation in a centralized department. In some respects the statistical work is duplicated just as many times as there are departments.

In the centralized organization an economy would generally be realized also in the number of employes required to perform the statistical work for the entire company. In order to give even an approach to efficient service each of the departmental operating units would have to be equipped to do all kinds of statistical computations. For instance, one or more computing machine operators would be essential to the efficient operation of each of these units. Thus eight operating departments of the company, having as many separate statistical units, would require at least eight computing machine operators. In a centralized statistical organization two or four efficient machine operators may be adequate to do

the work. Likewise, each of these units would require a supervisor in immediate control, while in the centralized organization one supervisor would probably be sufficient to supervise the entire operating force of the statistical organization.

6. *Mechanical Aids*—In the larger industrial organizations, where the use of the tabulating machine is required, the centralization of primary basic records is next to imperative. The waste in time and energy involved in the introduction of the tabulating machine in a decentralized organization is avoided. The use of such machines in the extreme type of decentralized organization may involve the punching, sorting and tabulating of the cards in each of the statistical units of the several departments. The folly of providing the punching, sorting, and tabulating machines separately for each of these units need not be emphasized. Certain of these machines are very expensive at the best, unless they can be kept continuously in use, and it is frequently a problem to keep the machines busy even in a centralized organization where the work of all departments is done. Especially is this true in the smaller business organizations.

7. *Comparative Departmental Figures*—One of the widest fields of service in the development of statistical data lies in serving the general officers with comparative departmental figures. When these figures are prepared separately by the several departments there is no assurance that the treatment of details will be identical and the final figures comparable. With the immediate supervision of the work broken up into as many disconnected parts as there are separate operating departments the probability of the figures being closely comparable is rather remote. And since the statistical units of the several operating departments, located as they are at a distance from the general offices, do not fully appreciate the importance of uniformity in the treatment of the data, the task of supervising them is an almost impossible one.

On the other hand, in the centralized statistical organization the minutest supervision can be given immediately to each and every operation involved in the development of the final figures. In other words, the development of comparable figures for the several departments necessitates that the detailed work be done centrally. Practical features alone

greatly hinder the working up of comparable figures for the several departments through a decentralized organization.

The Decentralized Organization.—On the other hand, there are certain important factors which argue strongly for the decentralized organization. The centralized organization may tend to relieve the operating departments of the burden of statistical work; it may be more economical; it may apparently be more efficient; it may be more conducive to the uniformity in the treatment of details and in the development of comparable departmental figures; and it may afford advantages from the standpoint of accuracy. In spite of these facts the decentralized organization offers certain advantages which frequently outweigh all of those listed above in favor of the centralized concept.

1. *Service to Department Heads*—We must not forget that it is important for the heads of departments to be familiar with the facts of their respective departments as well as for the general executives to have the facts of the company. The decentralized statistical organization affords a better opportunity to serve the department heads than does the centralized type. If the operating department is responsible for organizing and analyzing the facts of its own organization it will be more familiar with the facts and hence be better able to draw conclusions therefrom and establish policies and practices in accordance therewith. Furthermore, the department head can definitely control the statistical work which he requires and thus will be able to call for data without hesitation and require that they be furnished. He can also require immediate service and in the form that suits his own fancy. The very uniformity which is offered as an advantage in the centralized organization frequently becomes a disadvantage in that it usually results in the data being set up in the same way for all departments and consequently the individual department head cannot demand that he have figures organized to suit himself in the same way that he can when the organization is within his own department. From this point of view, the centralized organization is a violation of the principle that employes should not report to one man and serve another.
2. *Eliminates Duplication of Records*—A further advantage of the decentralized type is that it tends to eliminate the duplica-

tion of records which is required in the centralized organization. The operating departments must have their records for purposes of operations and the maintenance of a second set in the statistical department is definite duplication. The elimination of this may result in a substantial economy.

3. *Statistical Department Unfamiliar With the Facts*—In the centralized type of organization the figures are worked up entirely by the statistical organization without the inspection of the department or division heads. These figures presented to the executives may give the impression that something is seriously wrong and may lead to a request for an explanation from the department or division head. The fact may be that nothing is seriously wrong at all: there may be practical aspects to the situation which the figures do not reveal and which the statistical organization does not know. The evil effects of this procedure may be two-fold. In the first place it is apt to discredit the statistical organization in the minds of the department heads and, in the second place, requirements for explanations may become so great as to be burdensome to the department heads.

In the decentralized type of organization, any explanations required will be incorporated with the statement by the departmental statistical organization. The department employes are generally familiar with the fact and can see that the statements are accompanied with proper textual matter.

The Practical Form of Organization.—From the practical standpoint, satisfactory statistical service can be rendered through either type of organization. In strict theory, the centralized type is perhaps preferable but rarely if ever in practical business affairs can we make decisions on a strictly theoretical basis. In nearly every question of management that arises there are important practical considerations and the formulation of a statistical organization is no exception.

As a matter of fact, in practical work we generally have a combination of the centralized and decentralized type. External basic data generally and such internal basic data as are centrally recorded—perhaps payroll, labor turnover, sales or financial facts—as well as all company statements and departmental comparisons will be handled by the central organization. Other data,

particularly detailed operations facts related specifically to the several departments, may be handled separately by the respective departments. The facts which are worked up by the departments are then generally reported to the statistical organization, not in their original detail, but in a more or less classified and summarized form, from which company and comparative departmental figures may be prepared. What basic data will be maintained in the statistical organization and what in the other departments of the company will have to be decided in the light of the facts of the particular situation, the practical aspects of which will usually be the determining factors.

Frequently, if not usually, the statistical organization develops incidentally out of the interest and effort of some employe and the particular kind of data worked up and the way in which it is done are somewhat haphazard and unsystematic. The development of the statistical department of the Notaseme Hosiery Company, as related by the company's president, M. T. Fleisher, in the March 1922 issue of *System,* is typical. As indicated in this article the development from the standpoint of the data analyzed was as follows:

1. Data relating to shipment of finished goods.
2. When the data in No. 1 were well organized, the organization then turned to the analysis of the records of returned goods.
3. Data on handling of orders.
4. The preparation of production reports by individual workers in the mills—the proportion of perfect work and waste.

Finally, it must be emphasized that the statistical organization must be in the form that will best serve the particular company and it is unfortunate indeed for the practical statistician to undertake abruptly to bend the form of the statistical organization to suit some theoretical standard rather than to have it fit in with the organization and practices of the company. Figures 128 and 129 below are offered to illustrate the general forms of the two different types of organizations but these are merely suggestions to aid as a guide in the development of the organization and are in no sense set down as the only forms of successful statistical organization. It must always be borne in mind that the statistical

organization in a practical business concern may be of any form and size. It may be a single unit in the accounting department; it may be made up of several units in the different operating departments; or the statistical work may be done incidentally by the regular operating organization. We find the last method in operation for certain classes of data in practically every concern and it is un-

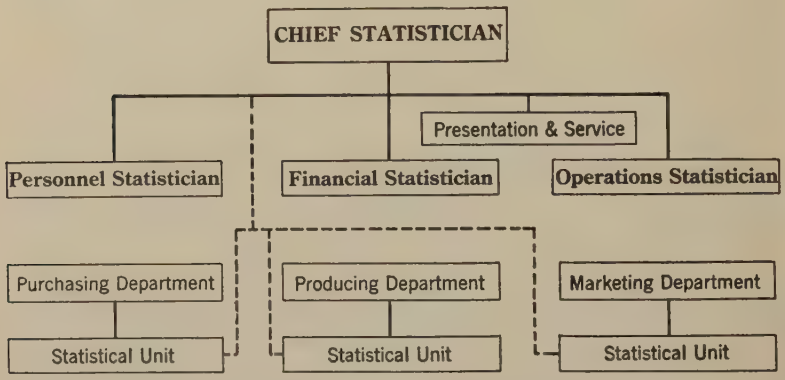


FIGURE 128.—An Organization Chart representing the Centralized Type of Statistical Organization with an executive statistician at the head and staff statisticians in charge of personnel, financial, and operations statistics, special inquiries, and presentation and service

Each of the main sub-divisions will have such technicians and expert statisticians as may be required. Such a complete organization will not be required for all types and sizes of concerns. The functions indicated, however, are generally applicable and the specific line-up will be selected in the light of the needs of the particular case. As noted on page 521 the chief statistician may report to the Operating Vice-president, the General Manager, the Comptroller or General Auditor, or some other official. Because of the similarity of records and character of work it seems preferable to have him report to the head of the accounting department.

doubtedly the preferable way of handling the statistics which are specifically applicable to the operation of a single department. What is here urged is merely that, other things being equal, the type of organization suggested herein is preferable in business concerns generally. But other things are rarely if ever equal and it is nonsense to say that such an ideal type of organization is indispensable to statistical service. Organizations differing widely therefrom are to be found in practical business concerns and the excellent statistical service which they are rendering is convincing on the point.

The Personnel.—The selection of proper personnel is one of the most important considerations in the development of a program of statistics. The accompanying organization charts suggest in general the type of personnel required. From these charts it will

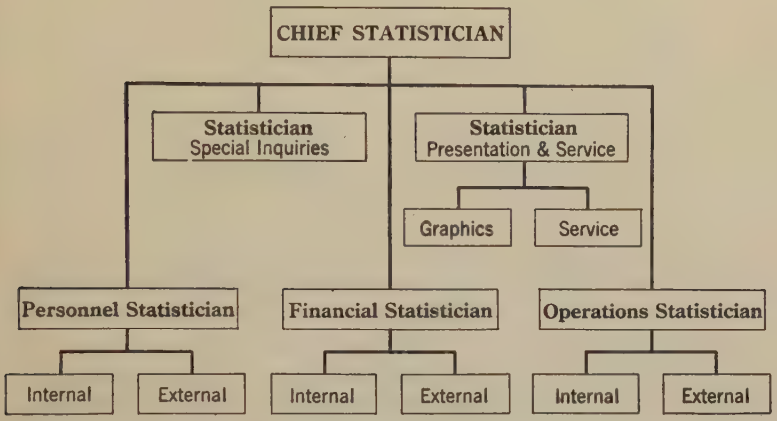


FIGURE 129.—Organization Chart Suggesting the General Form of the Decentralized Type of Statistical Organization

The extent to which decentralization is effected will vary greatly. Frequently if not generally the central organization will handle internal personnel and financial statistics and practically all external statistics. In any event the chief statistician will necessarily have staff statisticians of the types indicated for analytical and interpretative purposes if nothing else. He will also have certain graphics and computing work to do, depending on the extent to which such work is put on the statistical units of the company departments. The main function of the statistical units in the departments will be to maintain basic records and work up details but they may also do certain further analysis and some interpretative work.

be observed that the following broad types of employes are required:

1. Executive statisticians
2. Staff statisticians
3. Expert statisticians and technicians
4. Statistical clerks

THE EXECUTIVE STATISTICIAN.—The executive statistician is the one whose chief responsibility is the executive direction of the statistical organization. The following are among his principal qualifications:

1. *Administrative Ability*—The essential qualifications of the business executive are much the same whatever the type of organization he is directing and consequently a chief qualification of the executive statistician is administrative ability. He must *know what to do* and *how to do it.* In accomplishing this two-fold end he will depend much on the services of others and he must have the executive ability to bring their services to the accomplishment of the ends which he plans.
2. *An Appreciation of the Services of Statistics*—The executive statistician should be thoroughly familiar with the services which a statistical organization should render to a business concern. Obviously, this is essential to the proper visualization and development of the program.
3. *Disciplined Imagination*—The executive statistician should have disciplined imagination to visualize the complete program of statistics as he should like it ultimately to be. Somewhat along the lines suggested in Chapters II to VIII, he will work out on paper a program of statistics which he aims to build up and he will likewise outline a theoretical organization consistent with this program. He will generally keep both of these pictures to himself but he will use them as a guide in the constructive development of the program. He must also be able to foresee in broad outline, at least, and probably in detail, the outcome of any proposed line of statistical development and must be able to evaluate the results of such development in the management of the company.
4. *Familiarity With the Business*—Obviously, the ability to formulate the outline of detailed facts as suggested in Chapters II to VIII, presupposes a knowledge of the operations of the business. Furthermore, a knowledge of the problems of the business is necessary to a proper organization and analysis of these detailed facts. In brief, the executive statistician must be generally familiar with the operations of the company and the problems which are faced in its management if he is to be able to bring the force of his program of statistics satisfactorily to bear on these problems. He must have a practical knowledge of business management and of the fundamental operation of his own company.
5. *Statistical Technique*—While the executive statistician will have little personally to do with the technique of the statistical

work, he will find it necessary broadly to criticize this work and must be sufficiently familiar with technique to direct the work of his department, and particularly to see to it that the statistics which he releases are set up in the best form to serve the purpose intended.

6. *Practical Minded*—The executive statistician should be painfully practical. He must formulate his plans for the development of a program of statistics in accordance with the practical aspects of the company he is serving and be ever on the alert to guard against the development of a too theoretical program. There is a tendency in statistical work to become interested in merely the accumulation, analysis, and presentation of data—to forget that after all statistics are worth while only if they contribute to the primary aim of the business, namely: making profits. This is particularly true of the statistical experts and technicians on whom the executive statistician must largely depend for the actual development of his work and it requires the control and direction of a practical man to hold on to the straight and narrow path of practical service. At each step of the way, the question should be asked, "Will this be of practical benefit in the management of the business?"

THE STAFF STATISTICIAN.—The staff statistician may be designated as an assistant to the executive statistician, who is immediately in charge of the development of the statistical work in whole or in part. In the organization charts, Figures 128 and 129 above, provision has been made for staff statisticians in charge of personnel statistics, financial statistics, operations statistics, special inquiries, and presentation and service.

The qualifications of the staff statistician correspond in certain respects with those outlined above for the executive statistician. He also should have *administrative ability* that he may direct the activities of his corps of workers; *familiarity with the operations and processes of the company* that he may keep the development of his particular field of statistics closely within the requirements of the company; *practical minded* in order that he may always confine his statistical work to those developments which are of practical benefit to the company; and *disciplined imagination* in order that he may visualize the course of development of his phase

of the program and foresee the probable results of any proposed developments and evaluate them, before the work actually begins, in the light of the expense in time and money required in their development. All these are likewise qualifications of the staff statistician, but they need not be so prominent as in the executive statistician. The following are distinctive qualifications of the staff statistician:

1. *Analytical Powers*—If administrative ability is the chief qualification of the executive statistician, analytical ability is the chief qualification of the staff statistician. The burden of constructive statistical analysis rests primarily on him. It is he who is immediately responsible for the development of the statistical program through every step from the organization of basic records, to the analysis, the presentation, and the distribution of the facts. Hence, the weight of analysis falls directly on him and the expert assistants under his direction.
2. *Statistical Technique*—Likewise, the staff statistician will be immediately responsible for the technique of the statistical work. Here, also, he may be assisted by expert technicians but in many cases, particularly in smaller companies, he will have to bear this responsibility alone.

Statistical Experts.—These are employes who are specially equipped in the technique and philosophy of statistical method. The personnel, financial, operations, and graphic staff statisticians all will have assistants who are specialists each in his own field. These experts will give their entire time exclusively to the study of problems in their respective fields and will thereby give to the staff statisticians expert advice in the various phases of the statistical program.

The chief qualifications of the statistical expert are: (1) familiarty with the statistical methods required in the organization, analysis, and interpretation of the type of data for which he is responsible and, (2) a thorough knowledge of the important facts in his special field.

It should be noted that a corps of such experts as suggested by the organization charts shown in Figures 128 and 129 above, will be required only in the more elaborate statistical services.

Generally, in the smaller companies, one employe may be required to serve in this capacity for all types of data and in such cases the responsibilities for method and technique will be distributed more to those directing the organization.

Statistical Clerks.—This classification will include the office force of all classes: stenographers, typists, machine operators, draftsmen, and all kinds of clerks. While the selection of this class of employes is not generally so difficult as is the case with the executive and staff statistician, nevertheless it is highly important that they be carefully selected in the light of the special type of service they are to render. It is sometimes held that a "clerk is a clerk"—that statistical clerks can be readily obtained and easily trained—and that consequently little attention need be given to selecting the clerical force for a statistical organization. This is not the case. The statistical clerk has certain specific qualifications just as other employes of the organization.

The statistical clerks should have a liking and a sense for figures. It is true that much of the work to be performed by these employes is routine and that the statistical organization can be run with stenographers, typists, machine operators, draftsmen, and so forth, who are merely skilled in their respective activities. But it is also true that a much more satisfactory and efficient service can be rendered if these employes are naturally equipped to handle figures and can readily grasp the terminology and certain of the primary principles and methods of statistics. In fact, much will be gained if we can select clerks experienced along this line. The absence of such experience can be compensated for by giving training to them in the office either by preparing and distributing a statistical manual which contains the simple statistical processes and methods or by actually conducting classes for the purpose. The clerical force will of course pick up this information by experience if their natural qualifications are of the proper type.

The efficiency of the statistical organization will be greatly enhanced by having a high type of clerical help. It is much better to have fewer clerks and a higher quality rather than a greater number of those less efficient and less reliable.

The Number of Employes.—The kind and volume of statistical work to be done in various business concerns are not sufficiently standardized to enable us to indicate definitely the number of employes that will generally be required in a statistical department. This fact, however, does not confuse the situation. It would be of small advantage to know the personnel required in a completed program since in the proper method of procedure they will be engaged only as the program gradually develops.

The organization charts in Figures 128 and 129 above suggest the general character of the statistical organization and the type of personnel. As a matter of practice, however, we rarely find a completely rounded out program of this type. Hence, the number of employes will vary in all sorts of ways. In starting the program the first employee is, of course, the chief statistician with limited clerical and staff assistants. The problem of personnel is then taken care of simply by adding to the force as required. The program will usually be worked up gradually, starting with one type of statistics, say personnel or financial, and then gradually expanding the work as noted below. Hence employes can be taken on as the work requires.

Equipment.—The statistical organization should be amply supplied with computing and recording machines, the economies in the use of which are noted in Chapter XLV. The requirements along this line as in that of the personnel will vary so widely as to make it impossible to generalize. From the practical point of view it is necessary to keep constantly in touch with the various types of equipment and introduce them as their need appears with the development of the program of statistics. In this way the department will always be supplied with this equipment to the best advantage.

Cost of the Statistical Organization.—The cost of the statistical organization will vary not only as between companies of different sizes and types but also as between companies of the same size and type, because of difference in the type of organization and the volume and character of the work. Because of these variable factors it is impossible to generalize as to the cost of statistical work. Sufficient experience has been had in practical

business statistics, however, to justify the statement that the expense of a properly conducted statistical organization can be more than offset by the economies realized from it. M. T. Fleisher, president of the Notaseme Hosiery Company, writing in the March 1922 issue of *System,* says of the statistical department of his company:

"The department is saving us thousands of dollars without doubt, and enabling us to conduct our business more economically and efficiently. And it removes in a single stroke one of the most irksome problems of many an executive—the bother of getting facts to date when he wants them, and as he wants them."

From the practical point of view, the fundamental principle of business management will operate here as in other cases: the statistical organization must pay its way as it goes; and it has a place in the company only so long as it is financially profitable. In estimating its value, however, the broader services must be kept in mind. The statistical organization pays for itself primarily through furnishing systematic information relative to the operations and management of the business rather than through specific economies which result directly therefrom. Just how far it is wise to go in the development of details will depend on the type and size of business but we should ever bear in mind that there is a point of diminishing returns and this point should be determined in any given company. "How much can I afford to spend in checking up my processes?" is a live question to every business executive and one which can be answered only on the basis of the facts of the specific company. Greater detail will be generally required in the larger companies since the larger the organization the less the opportunity for the executive to keep directly in touch with the facts of his organization. It is becoming more and more imperative that the management of the business control its destiny through systematically collected, organized, analyzed, and interpreted facts and this is primarily the justification for the program of statistics regardless of the cost.

Conclusion.—This discussion has been designed merely to give broad suggestions as to the type of organization, the character and number of the personnel, the type and amount of equipment,

and the cost. The fundamental emphasis in this connection is that the program of statistics should be a gradual development. It is extremely unfortunate for the practical statistician to undertake to build up an elaborate organization at the outset of the statistical program: he should rather start the service and then expand the program and add personnel and equipment as the need justifies.

Specifically, the chief statistician should be selected and furnished with a secretary. He should then think his problem through and formulate a program and an outline of organization and the personnel required. At the start then he should select a certain kind of statistics, say personnel, and work these up and furnish the executives with analytical and interpretative statements concerning them. When this field has been fairly well developed he will then start another type of data, say financial, and then another, say operations, and proceed in the same manner. In this way the organization is always producing and the executives are being furnished with service whereby they can evaluate the work of the statistical organization.

The importance of this method of procedure in working up a program of statistics cannot be overemphasized. Many a noble effort in practical business statistics has come to grief for the simple reason that organization has been placed before service.

APPENDIX A

GLOSSARY OF STATISTICAL TERMS

ABSCISSA. A term used in graphics to designate the lines of the grid running parallel to the X-axis. The perpendicular distance of a point in the grid from the Y-axis—*i.e.*, its distance along the X-axis.

AMOUNT-OF-CHANGE CHART. See *Grid, Arithmetic.*

ANTILOGARITHM. An antilogarithm is the natural number which corresponds to a logarithm.

ARITHMETIC AVERAGE. See *Average.*

ARRAY. An array is an arrangement of the items of a group or series of numbers in order of size. The array may be in ascending order—from lowest to highest—or in descending order—from highest to lowest.

ASYMMETRY. See *Symmetrical Distribution.*

AVERAGE. A single numerical expression which stands for or represents a group of detailed data from which it has been derived. The average is sometimes referred to as the "type" or "central tendency."

Arithmetic Average. That amount which results from dividing the sum of a group of items by the number in the group. It is sometimes referred to as the "arithmetic mean" or simply as the "mean."

Geometric Average. The nth root of the product of the items in a group, "n" being the number of items in the group. That is, it is the result obtained by multiplying together the number of items in a group and then extracting a root of the product indicated by the number of items in the group.

Harmonic Average. The reciprocal of the arithmetic mean of the reciprocals of a series of statistical data. Thus the harmonic mean of 2, 4, and 6 is

$$\frac{1}{(1/2 + 1/4 + 1/6) \div 3}$$

Median. The value which divides the items of a statistical array into two equal groups. See *Quartiles, Deciles,* and *Percentiles.*

Mode. The mode "is the value occurring most frequently in a series of items and around which the other items are distributed most densely." —Zizek, "Statistical Averages," page 222.

Moving Averages. The averages of successive groups of a given number of items in a statistical series, each group after the first being formed by dropping the first item in the group immediately preceding and adding the next succeeding item in the series. Thus, in a three-period moving average of the following series the first group will consist of 10, 12,

and 15; the second of 12, 15, and 18; the third of 15, 18, and 21; and so on. The successive averages are usually centered at the mid-points of the groups.

10	
12	12.3
15	15.0
18	15.3
21	
27	
33	
36	

Periodic Average. A term applied to averages for corresponding time periods in different years, say, for January or for the first quarter of each year.

Progressive Mean. A weighted moving average, the coefficients of the binominal expansion being used as weights.

Quadratic Mean. The result obtained by taking the square root of the arithmetic average of the squares of the separate items. This method is used in the computation of the standard deviation.

Weighted Average. An average of a group of statistical data in the computation of which the relative importance of the separate items in the group has been altered through the effect of "weights" assigned to them. See Chapters XXXVI and XXXIX for methods of computing the weighted arithmetic, geometric, and harmonic averages. Extensive discussion may be found in Yule's "An Introduction to the Theory of Statistics," pages 220-225.

AVERAGE DEVIATION. See *Deviation.*

AXES. The two lines, one horizontal called the X-axis, and the other vertical called the Y-axis, which intersect at the point of origin, or the zero-point in a grid.

BOX. See *Table.*

CARTOGRAM. The type of statistical graph known as a statistical map—*i.e.*, a geographical map with statistical data represented in its different sections.

CASE. A term used to designate the statistical unit—a price, wage, day, age, ton or what not—which is the basis of the original observation.

CENTRAL TENDENCY. See *Average.*

CHARACTERISTIC. See *Logarithm.*

CHART. A term used in this text to designate a complete graphic presentation including the grid, scales, graph, title, key, and all other factors of the presentation. See graph.

CLASS FREQUENCY. See *Frequency Distribution.*

CLASS GROUPS. See *Frequency Distribution.*

CLASS INTERVAL. See *Frequency Distribution.*

CLASS LIMITS. See *Frequency Distribution.*

CLASS MARK. See *Frequency Distribution.*

CLASSIFICATION. "Classification in statistical methods consists in arranging data into groups according to their common characteristics."—Horace Secrist, "An Introduction to Statistical Methods," page 117.

COEFFICIENT OF CORRELATION. See *Correlation.*

CONTINUOUS SERIES. A statisical series made up of data of such character that the statistical unit may be made indefinitely small. Thus, the length of a bridge, for example, may be measured to the nearest yard, foot, inch, one-half inch, one-millionth inch, or smaller unit. By reducing the size of the unit the difference in the measurements becomes so small as to make the series approach continuity. See discontinuous series.

COORDINATES. See *Grid.*

CORRELATION. The tendency of two series of statistical data to vary together—*i.e.*, if the items of one fluctuate there is a tendency for corresponding items of the other to fluctuate in the same or opposite direction. If in the same direction, the correlation is positive or *direct;* if in the opposite direction, it is negative or *inverse.*

Coefficient. A coefficient of correlation is the numerical measurement of the correlation between series of data, the symbol for which is the letter "r."

Table. A correlation table is simply the combination of two series of data in tabular form, the one series of classes being arranged vertically to extend downward from a certain point and the other horizontally to extend to the right from the same point, so as to permit common entries for the separate pairs of classes.

DECILES. The values which divide the items of a statistical array into ten equal groups. See *Average—Median, Percentiles,* and *Quartiles.*

DEVIATION. A term used to designate the difference between the items in a statistical series and a standard, usually the average of the series.

Average Deviation (A. D.). It is the figure which results from dividing the sum (signs disregarded) of the deviations (from their average) of the several items of a series, by the number of items in the series. It is used as a measure of the dispersion of the items of a series of statistical data about their average, the mean, median, or mode.

Quartile Deviation (Q. D.). A measure of dispersion represented by one-half the distance between the first and third quartiles in a frequency series. This is reduced to the relative form or to the form of a coefficient by dividing the difference of the quartiles by their sum.

Range Deviation (R. D.). The difference between the greatest and smallest values in a statistical series.

Standard Deviation (S. D.). The square root of the arithmetic mean of the squares of the deviations of the several items in a statistical series from their average.

DISCONTINUOUS SERIES. A statistical series of which the minuteness of the unit of measurement is definitely limited by the character of the

unit of observation. Thus in counting the number of dots which show in separate throws of dice the unit could not be less than a single dot, or in counting the employes in a given company we could not get a real measurement less than one employe. It is also called *discrete* series. See *Continuous Series.*

DISCRETE SERIES. See *Discontinuous Series.*

DISPERSION. The variation of the items in a statistical series from the average of the series.

EXTRAPOLATION. The estimation (in tabular or graphic form) of the movement of a statistical series beyond the range of the actual data at hand. See *Interpolation.*

FIELD. See *Grid.*

FREQUENCY DISTRIBUTION (OR SERIES). A frequency distribution is an arrangement of a series of statistical values, called measurements, arrayed in ascending or descending order, there being associated with each measurement or group of measurements the number of times it occurs in the total observations, or its frequency. Table 5, page 385, is typical. The term is sometimes used in a broader sense to include all non-historical classification but it is generally confined to quantitative classification.

Class Frequency. The number of items falling within a given class in a frequency series is the class frequency of that class. It is also referred to simply as the frequency of the class.

Class Groups. The several classes of a frequency series.

Class Interval. The distance between the two limits of a class group in a frequency series.

Class Limits. The boundaries of a class group in a frequency series.

Class Mark. The mid-value of a class interval.

Frequency Curve. The graphical representation of a frequency series by means of the straight line or smoothed curve.

Frequency Polygon. The graph resulting from representing the class frequencies on the grid by means of rectangles erected on the horizontal base.

Frequency Table. The tabular representation of a frequency distribution.

GEOMETRIC AVERAGE. See *Average.*

GRAPH. The term "graph' is used in this text to designate the bar, curve, square, rectangle, cube, circle, picture or other device which is constructed on the grid to represent the data. It is distinguished from "chart" which is more comprehensive including the graph, the grid, the title, scales, and other parts of the graphic presentation. The term "graph," however, is used in statistical literature in the more comprehensive sense. The principal types are the bar, curve, area, volume, and picture graph.

Graph Designation or Caption. The descriptive wording associated with the graph to indicate the data which it represents.

GRID (FIELD). The grating-like base on which the statistical graph is constructed.

Arithmetic Grid. A grid on which equal vertical distances represent equal values and equal horizontal distances represent equal values. Other terms applied to this grid are plain, amount-of-change, difference, increment, rectilinear, and uniform.

Base Line. Those portions of the X-axis (horizontal base) and Y-axis (vertical base) which the grid comprehends or more loosely the bottom and left boundaries of the grid.

Coordinate Lines. The horizontal and vertical rulings in the grid. The "coordinates" of a point on the grid are its abscissa or the perpendicular distance from the Y-axis and its ordinate or the perpendicular distance from the X-axis.

Logarithmic Grid. A grid on which equal vertical distances represent equal ratios and equal horizontal distances represent equal ratios. The divisions on the grid are proportional to the logarithms of numbers. Other terms applied to this grid are ratio and rate-of-change.

Origin or Zero Point. The point where the X-axis and the Y-axis intersect. It is the starting point for all measurements on the grid. Measurements may extend in any direction from this point but in the most commonly used forms of grids they extend horizontally to the right and vertically upward.

Semi-Logarithmic Grid. A grid which has one set of rulings logarithmic and the other arithmetic. Usually the horizontal rulings (the vertical scale) is logarithmic.

HARMONIC AVERAGE. See *Average.*

HISTOGRAM. A term used in connection with the frequency graph, the precise meaning of which is not definitely standardized. To Professor W. I. King it includes the frequency graph in any form—rectangular, line-and-angle, or smoothed ("Elements of Statistical Method," pages 108-117). The "Manual of Charting" (Business School Series, page 36) and Professor Riegel, "Elements of Business Statistics" (pages 65-73), take a similar view. Mr. Karsten in his "Charts and Graphs" is somewhat indefinite in the meaning to give to the term. On page 205 he says, "A popular form of the curve chart is the 'staircase' curve, sometimes erroneously called a 'histogram.'" He reports it as synonymous with "historigram" on pages 220-221, where he states, "The historical curve has by some writers been called the 'histogram' or 'historigram' . . . . " And on page 329 he tacitly accepts the meaning which he condemns on page 205, as he remarks, "The staircase curve or histogram is always more accurate for period data . . . . " Finally, Professor Yule on page 84 of his "An Introduction to the Theory of Statistics" confines the meaning of the term to the rectangular form of the frequency graph and adds that the term was suggested by Professor Karl Pearson with this meaning: this meaning of the term is becoming

somewhat standardized. It is used in recent publications by Jerome in his "Statistical Method" (page 90); by Mills in his "Statistical Methods" (page 76); by Rugg, "Statistical Methods Applied to Education" (pages 88-93); and by Kelly "Statistical Method" (pages 9-15).

Historigram. The graphic representation of an historical series of statistical data.

Increment Chart. See *Grid—Arithmetic.*

Index Numbers. "Index numbers are series of numbers which measure and express the relative changes, as from time to time or from place to place, in the magnitude of statistical groups or aggregates of variables.—Rietz, "Handbook of Mathematical Statistics," page 181. An index series is sometimes called a "special purpose" series when it has been developed for a specific purpose or a "general purpose" series when it is comprehensive in the scope of data represented. Thus, an index constructed from the prices of the commodities of a single business concern would be a "special purpose" series and an index constructed from prices of different groups of commodities designed to reflect the general price movement would be a "general purpose" series.

Interpolation. The estimation of values lying between known values in a statistical series. Thus, in a grouped frequency series we interpolate to determine the modal or median value within the modal or median group; in a curve we interpolate to determine values of points which lie between known ordinates or abscissæ; and we interpolate to determine missing values within a statistical series.

Item. See *Case.*

Key. A separate scheme on a graphic chart consisting usually of specimens of the graphs which is designed to indicate the data which the several types or forms of graphs on the grid each represents.

Legend. See *Key.*

Logarithm. The logarithm of a given number is the exponent of the power to which the base of the system has to be raised to produce the given number. The Common System whose base is 10 is the one used most extensively. The logarithm of a number consists of two parts—the *characteristic* which is the part to the left of the decimal point and may be either positive or negative, and the *mantissa,* which is the part following the decimal point and is always positive. See anti-logarithm.

Logarithmic Grid. See *Grid.*

Mantissa. See *Logarithm.*

Mean Deviation. See *Deviation—Average.*

Median. See *Average.*

Mode. See *Average.*

Moving Average. See *Average.*

Moving Totals. The totals of successive groups of a given number of

items in a statistical series, each group after the first being formed by dropping the first item in the group immediately preceding and adding the next succeeding item in the series. In the illustration given under "moving average" above, the totals of the successive groups are "moving totals."

OGIVE. A curve representing the data of a cumulative frequency series.

ORDINATE. A term used in graphics to designate lines of the grid running parallel to the Y-axis. The perpendicular distance of a point in the grid from the X-axis—*i.e.*, its distance along the Y-axis is the ordinate of the point.

ORIGIN. See *Grid.*

PERCENTILES. The values which divide the items of a statistical array into a hundred equal groups. See *Average—Median, Deciles,* and *Quartiles.*

PERIODIC AVERAGE. See *Average.*

PERIODIGRAM. A curve which has more or less regular cyclical characteristics.

PICTOGRAM. A term in graphics the precise meaning of which is not definitely standardized. Professor Secrist defines the term as "the diagrams most commonly used to illustrate frequency and magnitude alone" ("An Introduction to Statistical Methods," page 163) and offers lines or bars, surfaces, volumes and picture graphs as typical. Professor King seems to hold a similar view ("Elements," page 92). The "Manual of Charting" (Business School Series) gives a narrower meaning to the term, including in it "any form or type of chart which has been disguised, distorted or converted in any way for the purpose of securing greater popular appeal" (page 71) and offers pictures of hogs, buildings, and lumps of rubber as illustrations. K. G. Karsten, in "Charts and Graphs" (pages 220-221 and 308), reports the term as being applied to frequency curves and points out the lack of its general acceptance. The term is not used extensively. No definite meaning attaches to it and it could well be dropped from statistical literature.

PLAIN RULING. See *Grid, Arithmetic.*

POPULATION. "'Universe' and 'population' are conventional expressions used in statistical terminology to designate the total or aggregate of items under consideration or from which a sample is taken. Thus, if one were taking a sample of Ford cars to determine the proportion of them equipped with speedometers, the universe would be the total number of Ford cars in use. This universe might be widened to include all automobiles or narrowed to include only 1924-model Fords owned by residents of a specified city."—H. Jerome, "Statistical Method," page 14*n.*

PROGRESSIVE MEAN. See *Average.*

QUARTILE DEVIATION. See *Deviation.*

QUARTILES. The values which divide the items of a statistical array into four equal groups. See *Median, Deciles and Percentiles.*

QUESTIONNAIRE (OR SCHEDULE). A written request from one person or agent to another for statistical information.

RANGE. See *Deviation.*

RATE-OF-CHANGE CHART. See *Grid, Logarithmic.*

RATIO CHART. See *Grid, Logarithmic.*

RELATIVE NUMBERS. Figures which represent the percentage relationship between the items of a statistical series and a selected value termed the base.

ROOT-MEAN-SQUARE—DEVIATION. See *Deviation—Standard.*

SAMPLING. The process whereby conclusions relative to a class or group of data are based on the analysis of a relatively small number of representative items. A statistical sample is a relatively small group of data which are representative of the entire group of the particular type or kind of data of which the sample is a part—*i.e.*, of the entire "population" or "universe."

SCALE. A term used in graphics to designate the graduations or calibrations on the base lines of the grid. Those on the horizontal base line are called the *horizontal scale* and those on the vertical base line the *vertical scale.*

Scale Designation or Caption. The descriptive wording associated with the scale to express the scale values in terms of the data represented by the graph.

Scale Unit. The portion of the base line which represents a unit of data.

SCHEDULE. See *Questionnaire.*

SCORE. See *Case.*

SEASONAL VARIATION. " . . . . the fluctuations in chronological data which follow the round of the seasons, exhibiting tendencies to be above normal in certain months of the year and below in others."—H. Jerome, "Statistical Method," page 234.

SECULAR TREND. A derived series of data which are designed to represent the general course of the series of original data, the fluctuations having been shortened or eliminated; or the approximation of such a series by short-cut computations or graphical methods.

SEMI-LOGARITHMIC RULING. See *Grid.*

SERIES. A general term used to designate a group of statistical data classified on some logical bases such as the historical classification based on the time element and the quantitative frequency classification based on quantitative measurements. It is sometimes narrowed to refer to historical classification only, the term group being applied to non-historical classification, but, this conception is not observed in general practice.

SKEWNESS. The lack of symmetry in a frequency distribution.

SMOOTHING. The process of ironing out from a statistical series or graph the irregular fluctuations due to "paucity of data or to unavoidable errors therein" in order to get an approximate representation of the general distribution or behavior of the data if the entire "population" were included in the analysis.

SPECIAL-PURPOSE INDEX NUMBERS. See *Index Numbers.*

STANDARD DEVIATION. See *Deviation.*

STATISTICAL UNIT. The unit in which a statistical observation is expressed. In counting wages the unit is a wage; in a population census, a single citizen; and in others a price, a ton, an interest rate, a sale, or any other item which is being counted.

STATISTICS. The term *statistics* is used in two distinct senses—first, to designate the body of methods and principles which are applied in the collection, presentation, and analysis of numerical facts, and, second, to designate the numerical facts themselves.

STEREOGRAM (OR STEREOGRAPH). A three dimensional graph designed to represent three related sets of data.

STUB. A term applied to the captions of the rows in a statistical table.

SYMMETRICAL DISTRIBUTION. A frequency distribution in which the class frequencies decrease symmetrically on either side of a central maximum.

Asymmetrical Distribution. A frequency series in which the distribution on opposite sides of the average or central maximum is unlike or non-symmetrical.

Moderately Asymmetrical Distribution. A frequency series in which the class-frequencies decrease with markedly greater rapidity on one side of the maximum than on the other.—G. U. Yule, "An Introduction to the Theory of Statistics," page 90.

TABLE. A statistical table is "an arrangement of figures in columns and rows so placed that the significance of the figures is determined by reference to two sets of explanatory phrases, *captions* at the head of the *columns* and *stubs* at the edge of the *rows.*"—H. Jerome "Statistical Method," page 29. See also page 246 *infra.*

TREND. See *Secular Trend.*

TYPE. See *Average.*

UNIVERSE. See *Population.*

VARIABILITY. Same as dispersion, which see.

VARIABLE. A term used in statistics to designate a characteristic whose numerical representation changes or *varies* in value. Thus we may say that ages, wages, prices, sales, imports, exports and similar items vary in value and hence are variables.

Variates. The individual values are called "variates."

Independent Variable. The variable in a statistical series the units of which are arbitrarily determined without regard for any other variable. It is the standard, so to speak, against which the data are measured or on the basis of which they are classified such as the time classification in a time series or the measurement classification in a frequency series.

Dependent Variable. The variable which is being measured against the standard or independent variable: it is a function, so to speak, of the independent variable.

WEIGHTED AVERAGE. See *Average.*

ZERO POINT. See *Grid.*

APPENDIX B

SUGGESTED STATISTICAL REPORTS AND STATEMENTS

The statistical organization will develop some facts as secondary basic data to be maintained in its own files, other facts to be distributed in the form of routine reports, and still others to be presented as special studies. The particular developments to be used in any given organization will be selected to suit the peculiar needs of the company and a general set cannot be prescribed which is best for all companies. Certain broad guides, however, can be set up which will be serviceable particularly to the beginner. Generally there will be need in all types and sizes of companies for reports and statements showing for all important groups of data:

I. Comparisons
   1. Between different periods of time.
   2. Between organizational units.
   3. Between groups or classes of facts such as current assets with current liabilities, sales with deliveries, rates of pay by occupational classes, and so forth.
   4. Of actual facts with established standards.

II. Changes in the Status of the Business
   1. Current—Running report.
   2. Long-time—Historical report.

III. Future Estimates
   1. Current or working estimates such as purchasing or production requirements. They are for relatively short periods, say, 3 to 6 months.
   2. Long-time estimates which partake of the nature of forecasts of the status or requirements of the business at some relatively long period in the future—a year or more. The most typical statement of this class is the annual company budget or provisional estimate.

Secondary basic records, routine reports, and special studies will be developed in all three of these groups. The factors determining just what shall be selected are so definitely tied up with the individual company being served that generalization in the matter is of little value. The records and reports given below are typical. Graphic charts presenting these and other important facts should also be developed.

PERSONNEL FACTS—NUMBER OF EMPLOYES

		Occupational Classes						Sex		Methods of Securing Employes						Base Pay per Week								Length of Service						Depart-ments			Labor Losses by Causes							Absences by Causes				Day Workers	Night Workers
		Office			Factory																																						
		Admin. & Supervis.	Clerical	Total	Admin. & Supervis.	Skilled Workers	Unskilled Workers	Total	Men	Women	Unsolicited	Through Employes	Agency	Newspaper Ad.	Under $12	$12 under $15	15 " 20	20 " 25	25 " 30	30 " 50	50 and over	Under 3 months	3 mo. under 6 mo.	6 mo. " 1 yr.	1 yr. " 2 yrs.	2 yrs. " 5 yrs.	5 yrs. and over	Purchasing	Producing	Marketing	Wages	Employm't Conditions	Better Position	Ill Health	Dismissal	Others	Total	Disability	Personal	Others	Total		
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42
1	Total																																									
2	Day Workers																																								
3	Night Workers																																								
4	Absences by Causes	Disability																																							
5		Personal																																							
6		Others																																							
7		Total																																							
8	Labor Losses by Causes	Wages																																							
9		Employm't Conditions																																							
10		Better Position																																							
11		Ill Health																																							
12		Dismissal																																							
13		Others																																							
14		Total																																							
15	Depart-ments	Purchasing																																							
16		Producing																																							
17		Marketing																																							
18	Length of Service	Under 3 Months																																							
19		3 mo. under 6 mo.																																							
20		6 mo. " 1 yr.																																							
21		1 yr. " 2 yrs.																																							
22		2 yrs. " 5 yrs.																																							
23		5 yrs. and over																																							

FIGURE 130.—Typical Secondary Basic Record to be Developed and Kept in the Statistical Organization

To economize space the last three rows have been omitted.

PERSONNEL FACTS—ANALYSIS OF LABOR LOSSES

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
		LENGTH OF SERVICE						DEPARTMENTS				AGE							PAY								SOURCE				
		Under 3 mo.	3 mo. under 6 mo.	6 mo. " 1 yr.	1 yr. " 2 yrs.	2 yrs. " 5 yrs.	5 yrs. and over	Purchasing	Producing	Marketing	General	Under 18	18 - 22	22 - 26	26 - 30	30 - 34	34 - 38	38 and over	$16 - $20	21 - 25	26 - 30	31 - 35	36 - 40	41 and over	Married	Single	Unsolicited	Through Employes	Agency	Advertising	
1	REASONS FOR LEAVING																													
2	Wages																													
3	Employm't Conditions																													
4	Better Position																													
5	Ill Health																													
6	Dismissal																													
7	Other																													
8	SOURCE																													
9	Unsolicited																													
10	Through Employes																													
11	Agency																													
12	Advertising																													
13	PAY																													
14	$16 - $20																													
15	21 - 25																													
16	26 - 30																													
	31 - 35																													

FIGURE 131.—Typical Secondary Basic Record to be Kept in the Statistical Organization

This record serves as a basis for special studies, routine reports, and replies to informal inquiries. To economize space the last two rows have been omitted.

Personnel Statistics

Total Authorized Payroll—By Departments

Departments	Total Authorized Payroll (Beginning of Month)						Additions to Payroll During Month						Deductions from Payroll During Month						Total Authorized Payroll (End of Month)					
	Men		Women		Total		Increases in Rates of Pay		Appointment of New Employes		Total		Decrease in Rates		Losses		Total		Men		Women		Total	
	Amount	% of Total Payroll	Amount	% of Total Payroll	Amount	% of Total Payroll	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women	Amount	% of Total Payroll	Amount	% of Total Payroll	Amount	% of Total Payroll
a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y

Figure 132.—Typical Running Report in the Personnel Field

It should be distributed as a routine report. Other classifications might be substituted for "Departments" in the first column, such as occupational classes and period of service groups. Separate forms might also be prepared for men and women, office and factory workers, and day and night workers.

PERSONNEL STATISTICS—BASE RATES OF PAY (WEEKLY BASIS)

Blank Company and Outside Concerns

BY OCCUPATIONAL CLASSES

OCCUPATIONAL CLASS	STARTING RATES					AVERAGE RATES					MAXIMUM RATES				
	Blank Company	GROUPS OF OUTSIDE CONCERNS				Blank Company	GROUPS OF OUTSIDE CONCERNS				Blank Company	GROUPS OF OUTSIDE CONCERNS			
		Group A	Group B	Group C	Group D		Group A	Group B	Group C	Group D		Group A	Group B	Group C	Group D
a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p

FIGURE 133.—A Typical Report Comparing Rates of Pay in the Company with Those Paid by Other Concerns in the Community

Such a report should be distributed quarterly, semi-annually, or annually, depending on the condition of the labor market.

Personnel Statistics

Average Rates of Pay, Overtime, Labor Turnover

		a	Entire Company			Purchasing Dept.			Producing Dept.			Sales Dept.		
			Men	Women	Total	Men	Women	Total	Men	Women	Total	Men	Women	Total
			b	c	d	e	f	g	h	i	j	k	l	m
1	Average Rate of Pay	Day Workers												
2		Night Workers												
3		Married												
4		Single												
5		Occup. Classes—Administrative												
6		Supervisory												
7		Clerical												
8		Foremen												
9		Skilled												
10		Unskilled												
11		Period of Service—Under 3 mo.												
12		3 mo. under 6 mo.												
13		6 mo. " 1 yr.												
14		1 yr. " 2 yrs.												
15		2 yrs. " 5 yrs.												
16		5 yrs. and over												
17	Overtime for Last Month	Amount Paid—Total												
18		Equivalent in regular time												
19		Hours Paid for*—Total												
20		Equivalent in regular time for one employe												
21														
22	Labor Turnover—Per Cent													

* Total hours worked multiplied by the overtime basis of payment. Thus if time-and-a-half is allowed for overtime and 20 hours is worked, the report would show 30 hours.

Figure 134.—Simple Form of Special Analytical Report in the Field of Personnel Statistics

Personnel Statistics

Analysis of Absence Data

	Cause of Absence		Entire Company		Dept. A		Dept. B		Dept. C		Dept. D		Dept. E	
			Men	Women	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women
	a		b	c	d	e	f	g	h	i	j	k	l	m
1	SICKNESS													
2	Number of Absences	Total												
3		% to all absent cases												
4		% of total employes who have absences												
5		Av. per employe in absence group												
6		% to total employes												
7		% to normal												
8	Time Lost	Total												
9		Av. per case (perhaps in days)												
10		Av. per absent employe												
11		Av. per employe (all employes)												
12		% to total time of absent employes												
13		% to total time of all employes												
14		% of the total absent time												
15	ACCIDENT													
16	Number of Cases	Total–within company												
17		–without company												
18		% to all absent cases												
19		Av. per employe in accident group												
20		Av. per employe in company												
21		% to normal												
22	Time Lost	Total												
23		Av. per case												
24		Av. per absent (accident) employe												
25		Av. per absent (all) employe												
26		Av. per employe (total)												
27		% to total time of absent (accident) employes												
28		% to total time of absent (all) employes												
29		% to total time of total employes												
30	Other Causes													
31	Total													

Figure 135.—Typical Form of Special Report in the Personnel Field

Personnel Statistics

Methods of Obtaining New Employes

	Method	Number of Employes Obtained			
		MEN		WOMEN	
		Married	Single	Married	Single
	a	b	c	d	e
1	Advertising—Total				
2	Newspapers				
3	Trade Papers				
4	Agencies—Total				
5	Private				
6	Institutional				
7	Societies–Fraternal				
8	–Technical				
9	School Authorities—Total				
10	Public				
11	Private				
12	Other Employes—Total				
13	Voluntary				
14	By request				
15	Customers				
16	Associated Companies—Total				
17	From their own force				
18	Recommended from outside				
19	Other Companies				
20	Prospect File of Applications				
21	Scouting				

Figure 136.—Typical Form of Special Report in the Personnel Field

Such a report should be prepared perhaps once or twice a year, depending on the extent to which new employes are being engaged.

Current Status of Business—Assets and Liabilities

	Accounting Classifications	End of Current Month			Same Date Last Year			End of Last Month			Dec. 31 Last Year		
		Absolute Amount	% to Total[1]	% to Standard	Absolute Amount	% Change[4]	% to Standard	Absolute Amount	% Change[4]	% to Standard	Absolute Amount	% Change[4]	% to Standard
	a	b	c	d	e	f	g	h	i	j	k	l	m
1	**ASSETS**												
2	CAPITAL ASSETS												
3	Intangible Capital												
4	Physical Property												
5	Investment Securities[2]												
6	Other Capital Assets												
7	CURRENT ASSETS												
8	Materials and Supplies												
9	Raw Materials												
10	Materials in Process												
11	Finished Products												
12	Marketable Securities[3]												
13	Receivables												
14	Cash												
15	DEFERRED ASSETS												
16	Sinking Fund Assets												
17	Prepayments												
18	Other Deferred Assets												
19	**LIABILITIES**												
20	CAPITAL LIABILITIES												
21	Capital Stock												
22	Funded Debt												
23	Other Capital Liabilities												
24	CURRENT LIABILITIES												
25	Payables												
26	Others												
27	OTHER LIABILITIES												
28	Reserves												
29	Miscellaneous												
30	SURPLUS												

[1] Each sub-classification will be entered as a percentage of its main heading—*e.g.*, "Capital Assets" against total "Assets" but "Intangible Capital" against "Capital Assets."
[2] Those required in maintaining the business intact such as stock of a controlled company.
[3] Those that can be disposed of without interfering with the complete operation of the business.
[4] Per cent change in comparison with current month.

FIGURE 137.—Typical Routine Report in the Financial Field

Such a report should be distributed monthly for assets and liabilities, and revenues and expenses.

Financial Statistics

Current Status of Business—Assets and Liabilities

Running Report Covering Current Period

Accounting Classifications	Beginning of Period			Added During the Period		Deducted During the Period		End of Period		
	Absolute Amount	% to Total*	% to Standard	Absolute Amount	% to Total*	Absolute Amount	% to Total*	Absolute Amount	% to Total*	% to Standard
a	b	c	d	e	f	g	h	i	j	k
ASSETS										
Capital Assets										
Intangible Capital										
Physical Property										
Investment Securities										
Other Capital Assets										
Current Assets										
Materials										
Raw Materials										
Materials in Process										
Finished Products										
Marketable Securities										
Receivables										
Cash										
Deferred Assets										
Sinking Fund Assets										
Prepayments										
Other Deferred Assets										
LIABILITIES										
Capital Liabilities										
Capital Stock										
Funded Debt										
Other Capital Liabilities										
Current Liabilities										
Payables										
Others										
Other Liabilities										
Reserves										
Miscellaneous										
Surplus										

* Each sub-classification will be entered as a percentage of its main heading—*e.g.*, "Capital Assets" against total "Assets" but "Intangible Capital" against "Capital Assets."

Figure 138.—Typical Running Report in Financial Field

Such a report should be distributed monthly or quarterly.

Financial Statistics
Current Status of Business
Comparison with Outside Companies—Assets and Liabilities

	Accounting Classifications	Blank Company		Outside Company #1		Outside Company #2		Outside Company #3	
		Amt.	% to Total[1]	Amt.	% to Total[1]	Amt.	% to Total[1]	Amt.	% to Total[1]
	a	b	c	d	e	f	g	h	i
1	**ASSETS**								
2	Capital Assets								
3	Intangible Capital								
4	Physical Property								
5	Investment Securities[2]								
6	Other Capital Assets								
7	Current Assets								
8	Materials and Supplies								
9	Marketable Securities[3]								
10	Receivables								
11	Cash								
12	Deferred Assets								
13	**LIABILITIES**								
14	Capital Liabilities								
15	Capital Stock								
16	Funded Debt								
17	Other Capital Liabilities								
18	Current Liabilities								
19	Payables								
20	Other Current Liabilities								
21	Other Liabilities								
22	Reserves								
23	Miscellaneous								
24	Surplus								

[1] Each sub-classification will be entered as a percentage of its main heading—*e.g.* "Capital Assets" against total "Assets" but "Intangible Capital" against "Capital Assets."

[2] Those required in maintaining the business intact, such as stock of a controlled company.

[3] Those that can be disposed of without interfering with the complete operation of the business.

Figure 139.—Typical Special Report in Financial Field

Financial Statistics

Source and Disposition of Revenue—Historical Statement

Months	Sources				Disposition				
	Main Product	By-Product	Miscellaneous	Total	Expenses	Interest	Dividends	Undivided Profits	Total
a	b	c	d	e	f	g	h	i	j
January									
February									
March									
April									
May									
June									
July									
August									
September									
October									
November									
December									

Figure 140.—Typical Historical Report in the Financial Field

Such a report should be prepared monthly or quarterly.

Financial Statistics—

Month	Purchasing Expense						Producing Expense										Marketing Expense								
	Salaries and Wages	Warehouse or Storage Expenses	Other Purchasing Expenses	Total			Direct		Indirect					Total			Salaries and Wages	Advertising	Credits and Collections	Delivery Expenses	Other Marketing Expenses	Total			
				Amount	% to Total	% to Stand'd	Labor	Material	Labor	Supplies	Factory Expenses	Maintenance and Repairs	Depreciation	Other	Amount	% to Total	% to Stand'd						Amount	% to Total	% to Stand'd
a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
Jan.																									
Feb.																									
Mch.																									
April																									
May																									
June																									
July																									
Aug.																									
Sept.																									
Oct.																									
Nov.																									
Dec.																									

Figure 142.—Secondary Basic Record to be Maintained

Financial Statistics

Revenues and Expenses—Actual Compared with Estimate

Classification	Estimated	Actual	
		Amount	% to Estimated
a	b	c	d
REVENUES—Total			
Main Products			
1			
2			
3			
Total			
By-Products			
1			
2			
3			
Total			
EXPENSES—Total			
Purchasing			
Producing			
Marketing			
Other			

Figure 141.—Typical Progress Report in the Financial Field

The classes in column "a" may be more detailed. Such a report should be prepared monthly. Similar reports should be prepared for purchasing, production, and sales.

Historical Data

Administrative Expense								Total Expense		Profits								
					Total					Gross Revenues		Gross Income		Net Income		Appropriated		
Salaries and Wages	Stationery and Printing	Employes' Service Activities	Finance and Accounts	Other Administrative Expenses	Amount	% to Total	% to Stand'd	Amount	% to Revenue	Amount	% to Plant Investment	Amount	% to Plant Investment	Amount	% to Plant Investment	Dividends	Other	Balance to Surplus
aa	bb	cc	dd	ee	ff	gg	hh	ii	jj	kk	ll	mm	nn	oo	pp	qq	rr	ss

in the Statistical Organization

Financial Statistics—Acquisition of Capital

Stocks and Bonds—Compared With Outside Securities

Month	STOCKS												BONDS					
	COMMON						PREFERRED						PRICE		INTEREST		YIELD	
	Price		Dividends		Yield		Price		Dividends		Yield							
	Company	Outside	Company	Outside	Company	Outside	Company	Outside	Company	Outside	Company	Outside	Company	Outside	Company	Outside	Company	Outside
aa	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r
Jan.																		
Feb.																		
Mch.																		
April																		
May																		
June																		
July																		
Aug.																		
Sept.																		
Oct.																		
Nov.																		
Dec.																		

Figure 143.—Secondary Basic Record to be Maintained in the Statistical Organization and Distributed Perhaps Quarterly

Price data may show the "high" and low" or preferably the average.

Operations Statistics—Purchasing

Current Status of Business

Materials and Supplies

Classes of Materials	End of Current Month									End of Last Month									Same Date Last Year									Dec. 31 Last Year								
	On Hand			On Purchase			Total			On Hand			On Purchase			Total			On Hand			On Purchase			Total			On Hand			On Purchase			Total		
	Amount	Day's Supply	% to Standard	Amount	Day's Supply	% to Standard	Amount	Day's Supply	% to Standard	Amount	Day's Supply	% to Standard	Amount	Day's Supply	% to Standard	Amount	Day's Supply	% to Standard	Amount	Day's Supply	% to Standard	Amount	Day's Supply	% to Standard	Amount	Day's Supply	% to Standard	Amount	Day's Supply	% to Standard	Amount	Day's Supply	% to Standard	Amount	Day's Supply	% to Standard
a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	aa	bb	cc	dd	ee	ff	gg	hh	ii	jj	kk
Raw Material																																			
1																																			
2																																			
3																																			
Total																																			
Completed Parts																																			
1																																			
2																																			
3																																			
Total																																			
Supplies																																			
1																																			
2																																			
3																																			
Total																																			

Figure 144.—Typical Routine Report in Field of Purchasing Statistics

Operations Statistics—Producing

Current Status of Business

Running Report

Items of Production	Stocks on Hand at Beginning of Month			Stocks Added During Month			Stocks Deducted During Month		Stocks on Hand at End of Month		
	Amount	Day's Supply	% to Standard	Amount	Day's Supply	% to Standard*	Amount	% to Standard	Amount	Day's Supply	% to Standard
a	b	c	d	e	f	g	h	i	j	k	l
Finished Product											
A											
B											
C											
Total											
Completed Parts											
A											
B											
C											
Total											

* Such as required rate of production.

Figure 145.—Typical Running Report in the Field of Production

It should be prepared monthly. A similar report may be prepared for materials and supplies.

Operations Statistics—Producing

Future Estimate

Progress Report

Items of Production	Present Status			Estimated Future Status					Current Production					
	Stocks on Hand	Day's Supply	% to Requirements	Estimated Production Next 6 Months	Estimated Shipments Next 6 Months	Status of Business 6 Months Hence			Average for Last 4 Weeks		Average for Current Week		Required Average Next 6 Months	
						Stocks on Hand *	Day's Supply	% to Requirements	Amount	% to Schedule	Amount	% to Schedule	Amount	% to Schedule
a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
Finished Products														
1														
2														
3														
Total														
Completed Parts														
1														
2														
3														
Total														

* Columns b+e−f.

Figure 146.—Typical Progress Report in Field of Production Statistics

Operations Statistics—Sales (Current Status of Business)

	Cancellations	Deliveries	SALES																												
			Company's or Branch Office			Classes of Products			Cost			Price			Customers					Salesmen			Days of Week								
			A	B	C	1	2	3	1	2	3	1	2	3	1	2	3	Total Old Customers	Total New Customers	1	2	3	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday			
a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	aa	bb	cc			
Dist.—1 2 3																															
Days of week—M. T. W. Th. F. S.																															
Salesmen—1 2 3																															
Customers—1 2 3 Total—Old New																															
Selling Price—1 2 3																															
Cost—1 2 3																															
Classes of Products—1 2 3																															
Company's or Branch Offices—A B C																															

References

Franklin, B. A. Cost Reports for Executives, 1913, Ch VII.
Haskell, Allan C. Graphic Charts in Business, 1924, Chs. XXI-XXVII, XXIX-XXXIII.
Manual of Charting (Business School Series), 1923, Chs. VI-IX.
Savage, W. A. Graphic Analysis for Executives, 1924, Chs. II-VI.
Smith, W. H. Graphic Statistics in Management, 1924, Part IV.

APPENDIX C

PERTINENT FACTS RELATIVE TO A SELECTED GROUP OF AMERICAN INDEX NUMBERS

Character of Index	Prepared by	Published in	Period Covered	Frequency	Base Period	Number of Items	REMARKS
a	b	c	d	e	f	g	h
Wholesale Prices—U. S.	U. S. Bureau, Labor Statistics	Monthly Labor Review*	1913–date	Monthly	1913	404 (9 Groups)	1890–1921—By years in Review—Sept. 1922 1913–1922—By months in Review—July 1922
Wholesale Prices—U. S. and Foreign Countries—Int'l Price Index	Federal Reserve Board	Federal Reserve Bulletin*	1919–date	Monthly	1913	101	Shown separately by Stage of Manufacture, Origin, and Export Goods Description of Index—Bulletins May 1920, 1922, June 1920, 1921
Wholesale Prices—U. S.	R. G. Dun & Co.	Dun's Review*	1860–date	Monthly	None**	300 (7 Groups)	First Weekly Issue of Month Description in U. S. Bureau Labor Statistics—Bulletin 284
Wholesale Prices—U. S.	Bradstreet Co.	Bradstreet's Journal*	1892–1898 1899–date	Quarterly Monthly	None**	96 (13 Groups)	Second Weekly Issue of Month Description in U. S. Bureau Labor Statistics—Bulletin 284
Retail Food Prices—U. S.	U. S. Bureau, Labor Statistics	Monthly Labor Review*	1907–date	Monthly	1913	43 (22 Groups)	From Jan. 1913–Dec. 1920—22 Articles Description—Review—Mar. 1921; Early Figures—Survey July 1922
Retail Food Prices—U. S. and Foreign Countries	Federal Reserve Board	Federal Reserve Bulletin	Oct. 1923–date	Monthly	Pre-war=100		Information available in Board's Office, Data collected from 51 cities
Farm Prices—U. S.	U. S. Dept. of Agriculture	Crops and Markets*	1913–date	Monthly	1909–14 Avge.	31 (6 Groups)	Description in Crops & Markets—Aug. 1924—Supplement
Cost of Living—U. S.	U. S. Bureau, Labor Statistics	Monthly Labor Review*	Dec. 1914–date	Quarterly	1913 Avge.	6 Groups	Methods explained in the Review. Index for 32 cities separately and combined
Cost of Living—U. S.	National Industrial Conference Board	Press Release* 21st of Month	1914–date	Monthly	July 1914		Description in Survey, Feb. 1925 and Nat'l Ind. Conf. Bd. Research Report #63
Cost of Living—Mass. Compared with Foreign Countries	Federal Reserve Board	Federal Reserve Bulletin*	Oct. 1923–date	Monthly	Pre-war=100		Information available on Board's Office
Production in Basic Industries—U. S.	Federal Reserve Board	Federal Reserve Bulletin	1913–date	Monthly	1919	22 (11 Groups)	Corrected for Seasonal Element, Description in Bulletins: March 1922, March 1924, Dec. 1922
Production in Basic Industries —U. S.	U. S. Dept. of Commerce	Survey of Current Business	1919–date	Monthly	1919	64—Manufacturing 57 Raw Mat. (4 Groups)	Description in Survey—Jan. 1923, 1924, 1925 Description in Survey—May–Sept. 1922

* See page 570 for note. ** For the series which appears in the Survey of Current Business, 1913 is used as base.

PERTINENT FACTS RELATIVE TO A SELECTED GROUP OF AMERICAN INDEX NUMBERS (Contd.)

CHARACTER OF INDEX	PREPARED BY	PUBLISHED IN	PERIOD COVERED	FREQUENCY	BASE PERIOD	NUMBER OF ITEMS	REMARKS
a	b	c	d	e	f	g	h
Wholesale Trade—U. S. Avge. Mo. Sales	Federal Reserve Board	Federal Reserve Bulletin	Oct. 1923–date	Monthly	1919	6 Groups	Monthly sales by lines
Retail Trade—U. S. Avge. Mo. Sales	Federal Reserve Board	Federal Reserve Bulletin	Oct. 1923–date	Monthly	1919	9 Groups	Monthly Sales by reporting lines with and without seasonal adjustment
Concrete Building Costs—U. S.	Aberthaw Const. Co.	Constr. Trade Papers*	1921–date	Monthly	1914		Cost of Material, Labor, Plant and Overhead Items in constructing a standard concrete factory building
Construction Cost—U. S.	McGraw-Hill Co., Inc.	Eng'r'g News Record*	1913–date	Monthly	1913	4 Groups	First Weekly Issue of Month Description in Survey—Feb. 1925
Construction Volume—U. S.	McGraw-Hill Co., Inc.	Eng'r'g News Record*	1920–date	Monthly	1913		Compiled on basis of contracts let
Average Weekly Earnings—U. S.	National Industrial Conference Board	Press Release* 21st of Month	1920–date	Monthly	1914	1678 Mfg. Plants	23 Industries; Description in Survey—Feb. 1925
Average Weekly Earnings—New York State	N. Y. State Dept. of Labor	Industrial Bulletin*	June 1914 –date	Monthly	June 1914	1648 Firms	Description in N. Y. State Dept. of Labor Special Bulletin #85, July 1917
Union Wage Scales—U. S.	U. S. Bureau, Labor Statistics	Annual Bulletin	1907–date	Annually	1913		Published as of May each year
Hours of Labor—U. S.	National Industrial Conference Board	Press Release* 21st of Month	1920 date	Monthly	1914	1678 Mfg. Plants	Description in Survey—Feb. 1925
Employment—Manufacturing Ind.—U. S.	U. S. Bureau Labor Statistics	Monthly Labor Review*	June 1914 –date	Monthly	1923		Description in Review—April 1924
Number of Employes—New York State	N. Y. State Dept. of Labor	Industrial Bulletin*	June 1914 –date	Monthly	June 1914	1648 Firms	Description in N. Y. State Dept. of Labor Special Bulletin #85, July 1917
Employment—Manufacturing Ind.—U. S.	Federal Reserve Board	Federal Reserve Bulletin	1919–date	Monthly	1919	33 Industries	Description and Early Figures—Bulletin, Dec. 1923
Business Summary—U. S.	U. S. Dept. of Commerce	Survey of Current Business	1919–date	Monthly	1919	22 (10 Groups)	Description and Early Figures—Survey, April, 1924
Business Indicators—U. S.	U. S. Dept. of Commerce	Survey of Current Business	1920–date	Monthly	{1st Part—1913 {2nd Part—1919	{28 (8 Groups) {9 (4 Groups)	Table divided into two parts Description in Survey—Feb. 1925
Industrial Activity—U. S.	Federal Reserve Board	Federal Reserve Bulletin	1919–date	Monthly	1919	26 (3 Groups)	Description and Early Figures—Bulletin, Mar. 1924

* Monthly figures appear also in current issues of the Survey of Current Business. Monthly figures for the past two years and annual comparison, where available back to 1913, appear in detailed tables published twice a year in Survey of Current Business.

APPENDIX D

TO FIND THE LOGARITHM

First determine the characteristic in accordance with the following rules:

1. The characteristic of a number *greater than unity* is positive and one less than the number of digits preceding the decimal point. Thus, the characteristics of the logarithms of 3406, 340.6, 34.06, and 3.406 are respectively 3, 2, 1, and 0.
2. The characteristic of a number *less than unity* is negative and one more than the number of ciphers immediately after the decimal point. Thus, the characteristics of the logarithms of .3406, .03406, .003406, and .0003406 are respectively -1, -2, -3, and -4. The negative characteristic is usually indicated by writing the minus sign over it to indicate that it alone is negative and not the mantissa. The logarithm of .0003406, for example, is written $\bar{4}.5322$.

Second, determine the mantissa directly from the table. In doing so, the position of the decimal point in the number is disregarded—*e.g.*, .1256, 1.256, 12.56, and 1256 are all considered as 1256. To obtain the mantissa of any given number proceed as follows:

3. Find the first two digits of the given number in column "N" of the table, the third digit in the series of figures given in group A at the top of the table, and the fourth digit in the series given in group B.
4. In the horizontal row headed by the first two digits of the number and the column headed by the third digit, the mantissa of a number made up of the first three digits of the given number taken in order will be found. Add to the right-hand figures of this mantissa the number given in the same horizontal row and in the column headed by the fourth digit. This is the required mantissa.

Example: Find the logarithm of 2534. In the row headed by 25 in column "N" and in the column headed by 3 in series A, will be found .4031 which is the mantissa of 253. In the same row and in the column headed by 4 in series B we find 7 which added to .4031 gives .4038, the mantissa of 2534. In accordance with rule 1 above the characteristic is 3 and hence the complete logarithm is 3.4038.

The series of columns in group B represent differences arrived at by interpolation according to the following rule:[1]

Find the tabular mantissa of the first three significant figures of the given number; subtract this mantissa from the next greater tabular mantissa; multiply the difference thus found by the remaining figures of the number, as a decimal; this will give the differences shown in the table.

[1] E. A. Bowser, Logarithmic Tables.

Thus in the number 2534 given above, the mantissa for 253 is .4031 and for 254, .4048. The difference between these two is 17 and this multiplied by .4, the final digit in the number, gives 6.8 or 7, as the difference to be added. The numbers listed in group B are these differences worked out for convenience. This method of interpolation will be followed in using tables where the differences are not given in detail.

TO FIND THE ANTILOGARITHMS

The number corresponding to a given logarithm may be found in similar manner from the table of antilogarithms. To find the antilogarithm of 3.4038, we locate .40 in column "L," 3 in group A and thus determine 2529 as the antilogarithm of .403. To this we add the difference, 5, noted in the column headed by 8 in group B which gives 2534 as a result. The differences in this table are also the result of interpolation.

Four-Place Logarithms—Common System

N	A										B		
	0	1	2	3	4	5	6	7	8	9	1 2 3	4 5 6	7 8 9
10	0000	.0043	.0086	.0128	.0170	.0212	.0253	.0294	.0334	.0374	4 8 12	17 21 25	29 33 37
11	.0414	.0453	.0492	.0531	.0569	.0607	.0645	.0682	.0719	.0755	4 8 11	15 19 23	26 30 34
12	.0792	.0828	.0864	.0899	.0934	.0969	.1004	.1038	.1072	.1106	3 7 10	14 17 21	24 28 31
13	.1139	.1173	.1206	.1239	.1271	.1303	.1335	.1367	.1399	.1430	3 6 10	13 16 19	23 26 29
14	.1461	.1492	.1523	.1553	.1584	.1614	.1644	.1673	.1703	.1732	3 6 9	12 15 18	21 24 27
15	.1761	.1790	.1818	.1847	.1875	.1903	.1931	.1959	.1987	.2014	3 6 8	11 14 17	20 22 25
16	.2041	.2068	.2095	.2122	.2148	.2175	.2201	.2227	.2253	.2279	3 5 8	11 13 16	18 21 24
17	.2304	.2330	.2355	.2380	.2405	.2430	.2455	.2480	.2504	.2529	2 5 7	10 12 15	17 20 22
18	.2553	.2577	.2601	.2625	.2648	.2672	.2695	.2718	.2742	.2765	2 5 7	9 12 14	16 19 21
19	.2788	.2810	.2833	.2856	.2878	.2900	.2923	.2945	.2967	.2989	2 4 7	9 11 13	16 18 20
20	.3010	.3032	.3054	.3075	.3096	.3118	.3139	.3160	.3181	.3201	2 4 6	8 11 13	15 17 19
21	.3222	.3243	.3263	.3284	.3304	.3324	.3345	.3365	.3385	.3404	2 4 6	8 10 12	14 16 18
22	.3424	.3444	.3464	.3483	.3502	.3522	.3541	.3560	.3579	.3598	2 4 6	8 10 12	14 15 17
23	.3617	.3636	.3655	.3674	.3692	.3711	.3729	.3747	.3766	.3784	2 4 6	7 9 11	13 15 17
24	.3802	.3820	.3838	.3856	.3874	.3892	.3909	.3927	.3945	.3962	2 4 5	7 9 11	12 14 16
25	.3979	.3997	.4014	.4031	.4048	.4065	.4082	.4099	.4116	.4133	2 3 5	7 9 10	12 14 15
26	.4150	.4166	.4183	.4200	.4216	.4232	.4249	.4265	.4281	.4298	2 3 5	7 8 10	11 13 15
27	.4314	.4330	.4346	.4362	.4378	.4393	.4409	.4425	.4440	.4456	2 3 5	6 8 9	11 13 14
28	.4472	.4487	.4502	.4518	.4533	.4548	.4564	.4579	.4594	.4609	2 3 5	6 8 9	11 12 14
29	.4624	.4639	.4654	.4669	.4683	.4698	.4713	.4728	.4742	.4757	1 3 4	6 7 9	10 12 13
30	.4771	.4786	.4800	.4814	.4829	.4843	.4857	.4871	.4886	.4900	1 3 4	6 7 9	10 11 13
31	.4914	.4928	.4942	.4955	.4969	.4983	.4997	.5011	.5024	.5038	1 3 4	6 7 8	10 11 12
32	.5051	.5065	.5079	.5092	.5105	.5119	.5132	.5145	.5159	.5172	1 3 4	5 7 8	9 11 12
33	.5185	.5198	.5211	.5224	.5237	.5250	.5263	.5276	.5289	.5302	1 3 4	5 6 8	9 10 12
34	.5315	.5328	.5340	.5353	.5366	.5378	.5391	.5403	.5416	.5428	1 3 4	5 6 8	9 10 11
35	.5441	.5453	.5465	.5478	.5490	.5502	.5514	.5527	.5539	.5551	1 2 4	5 6 7	9 10 11
36	.5563	.5575	.5587	.5599	.5611	.5623	.5635	.5647	.5658	.5670	1 2 4	5 6 7	8 10 11
37	.5682	.5694	.5705	.5717	.5729	.5740	.5752	.5763	.5775	.5786	1 2 3	5 6 7	8 9 10
38	.5798	.5809	.5821	.5832	.5843	.5855	.5866	.5877	.5888	.5899	1 2 3	5 6 7	8 9 10
39	.5911	.5922	.5933	.5944	.5955	.5966	.5977	.5988	.5999	.6010	1 2 3	4 5 7	8 9 10
40	.6021	.6031	.6042	.6053	.6064	.6075	.6085	.6096	.6107	.6117	1 2 3	4 5 6	8 9 10
41	.6128	.6138	.6149	.6160	.6170	.6180	.6191	.6201	.6212	.6222	1 2 3	4 5 6	7 8 9
42	.6232	.6243	.6253	.6263	.6274	.6284	.6294	.6304	.6314	.6325	1 2 3	4 5 6	7 8 9
43	.6335	.6345	.6355	.6365	.6375	.6385	.6395	.6405	.6415	.6425	1 2 3	4 5 6	7 8 9
44	.6435	.6444	.6454	.6464	.6474	.6484	.6493	.6503	.6513	.6522	1 2 3	4 5 6	7 8 9
45	.6532	.6542	.6551	.6561	.6571	.6580	.6590	.6599	.6609	.6618	1 2 3	4 5 6	7 8 9
46	.6628	.6637	.6646	.6656	.6665	.6675	.6684	.6693	.6702	.6712	1 2 3	4 5 6	7 7 8
47	.6721	.6730	.6739	.6749	.6758	.6767	.6776	.6785	.6794	.6803	1 2 3	4 5 5	6 7 8
48	.6812	.6821	.6830	.6839	.6848	.6857	.6866	.6875	.6884	.6893	1 2 3	4 4 5	6 7 8
49	.6902	.6911	.6920	.6928	.6937	.6946	.6955	.6964	.6972	.6981	1 2 3	4 4 5	6 7 8
50	.6990	.6998	.7007	.7016	.7024	7033	.7042	.7050	.7059	.7067	1 2 3	3 4 5	6 7 8
51	.7076	.7084	.7093	.7101	.7110	.7118	.7126	.7135	.7143	.7152	1 2 3	3 4 5	6 7 8
52	.7160	.7168	.7177	.7185	.7193	.7202	.7210	.7218	.7226	.7235	1 2 2	3 4 5	6 7 7
53	.7243	.7251	.7259	.7267	.7275	.7284	.7292	.7300	.7308	.7316	1 2 2	3 4 5	6 6 7

FOUR-PLACE LOGARITHMS—COMMON SYSTEM (Contd.)

N	A										B		
	0	1	2	3	4	5	6	7	8	9	1 2 3	4 5 6	7 8 9
54	.7324	.7332	.7340	.7348	.7356	.7364	.7372	.7380	.7388	.7396	1 2 2	3 4 5	6 6 7
55	.7404	.7412	.7419	.7427	.7435	.7443	.7451	.7459	.7466	.7474	1 2 2	3 4 5	5 6 7
56	.7482	.7490	.7497	.7505	.7513	.7520	.7528	.7536	.7543	.7551	1 2 2	3 4 5	5 6 7
57	.7559	.7566	.7574	.7582	.7589	.7597	.7604	.7612	.7619	.7627	1 2 2	3 4 5	5 6 7
58	.7634	.7642	.7649	.7657	.7664	.7672	.7679	.7686	.7694	.7701	1 1 2	3 4 4	5 6 7
59	.7709	.7716	.7723	.7731	.7738	.7745	.7752	.7760	.7767	.7774	1 1 2	3 4 4	5 6 7
60	.7782	.7789	.7796	.7803	.7810	.7818	.7825	.7832	.7839	.7846	1 1 2	3 4 4	5 6 6
61	.7853	.7860	.7868	.7875	.7882	.7889	.7896	.7903	.7910	.7917	1 1 2	3 4 4	5 6 6
62	.7924	.7931	.7938	.7945	.7952	.7959	.7966	.7973	.7980	.7987	1 1 2	3 3 4	5 6 6
63	.7993	.8000	.8007	.8014	.8021	.8028	.8035	.8041	.8048	.8055	1 1 2	3 3 4	5 5 6
64	.8062	.8069	.8075	.8082	.8089	.8096	.8102	.8109	.8116	.8122	1 1 2	3 3 4	5 5 6
65	.8129	.8136	.8142	.8149	.8156	.8162	.8169	.8176	.8182	.8189	1 1 2	3 3 4	5 5 6
66	.8195	.8202	.8209	.8215	.8222	.8228	.8235	.8241	.8248	.8254	1 1 2	3 3 4	5 5 6
67	.8261	.8267	.8274	.8280	.8287	.8293	.8299	.8306	.8312	.8319	1 1 2	3 3 4	5 5 6
68	.8325	.8331	.8338	.8344	.8351	.8357	.8363	.8370	.8376	.8382	1 1 2	3 3 4	4 5 6
69	.8388	.8395	.8401	.8407	.8414	.8420	.8426	.8432	.8439	.8445	1 1 2	2 3 4	4 5 6
70	.8451	.8457	.8463	.8470	.8476	.8482	.8488	.8494	.8500	.8506	1 1 2	2 3 4	4 5 6
71	.8513	.8519	.8525	.8531	.8537	.8543	.8549	.8555	.8561	.8567	1 1 2	2 3 4	4 5 5
72	.8573	.8579	.8585	.8591	.8597	.8603	.8609	.8615	.8621	.8627	1 1 2	2 3 4	4 5 5
73	.8633	.8639	.8645	.8651	.8657	.8663	.8669	.8675	.8681	.8686	1 1 2	2 3 4	4 5 5
74	.8692	.8698	.8704	.8710	.8716	.8722	.8727	.8733	.8739	.8745	1 1 2	2 3 4	4 5 5
75	.8751	.8756	.8762	.8768	.8774	.8779	.8785	.8791	.8797	.8802	1 1 2	2 3 3	4 5 5
76	.8808	.8814	.8820	.8825	.8831	.8837	.8842	.8848	.8854	.8859	1 1 2	2 3 3	4 5 5
77	.8865	.8871	.8876	.8882	.8887	.8893	.8899	.8904	.8910	.8915	1 1 2	2 3 3	4 4 5
78	.8921	.8927	.8932	.8938	.8943	.8949	.8954	.8960	.8965	.8971	1 1 2	2 3 3	4 4 5
79	.8976	.8982	.8987	.8993	.8998	.9004	.9009	.9015	.9020	.9025	1 1 2	2 3 3	4 4 5
80	.9031	.9036	.9042	.9047	.9053	.9058	.9063	.9069	.9074	.9079	1 1 2	2 3 3	4 4 5
81	.9085	.9090	.9096	.9101	.9106	.9112	.9117	.9122	.9128	.9133	1 1 2	2 3 3	4 4 5
82	.9138	.9143	.9149	.9154	.9159	.9165	.9170	.9175	.9180	.9186	1 1 2	2 3 3	4 4 5
83	.9191	.9196	.9201	.9206	.9212	.9217	.9222	.9227	.9232	.9238	1 1 2	2 3 3	4 4 5
84	.9243	.9248	.9253	.9258	.9263	.9269	.9274	.9279	.9284	.9289	1 1 2	2 3 3	4 4 5
85	.9294	.9299	.9304	.9309	.9315	.9320	.9325	.9330	.9335	.9340	1 1 2	2 3 3	4 4 5
86	.9345	.9350	.9355	.9360	.9365	.9370	.9375	.9380	.9385	.9390	1 1 2	2 3 3	4 4 5
87	.9395	.9400	.9405	.9410	.9415	.9420	.9425	.9430	.9435	.9440	0 1 1	2 2 3	3 4 4
88	.9445	.9450	.9455	.9460	.9465	.9469	.9474	.9479	.9484	.9489	0 1 1	2 2 3	3 4 4
89	.9494	.9499	.9504	.9509	.9513	.9518	.9523	.9528	.9533	.9538	0 1 1	2 2 3	3 4 4
90	.9542	.9547	.9552	.9557	.9562	.9566	.9571	.9576	.9581	.9586	0 1 1	2 2 3	3 4 4
91	.9590	.9595	.9600	.9605	.9609	.9614	.9619	.9624	.9628	.9633	0 1 1	2 2 3	3 4 4
92	.9638	.9643	.9647	.9652	.9657	.9661	.9666	.9671	.9675	.9680	0 1 1	2 2 3	3 4 4
93	.9685	.9689	.9694	.9699	.9703	.9708	.9713	.9717	.9722	.9727	0 1 1	2 2 3	3 4 4
94	.9731	.9736	.9741	.9745	.9750	.9754	.9759	.9763	.9768	.9773	0 1 1	2 2 3	3 4 4
95	.9777	.9782	.9786	.9791	.9795	.9800	.9805	.9809	.9814	.9818	0 1 1	2 2 3	3 4 4
96	.9823	.9827	.9832	.9836	.9841	.9845	.9850	.9854	.9859	.9863	0 1 1	2 2 3	3 4 4
97	.9868	.9872	.9877	.9881	.9886	.9890	.9894	.9899	.9903	.9908	0 1 1	2 2 3	3 4 4
98	.9912	.9917	.9921	.9926	.9930	.9934	.9939	.9943	.9948	.9952	0 1 1	2 2 3	3 4 4
99	.9956	.9961	.9965	.9969	.9974	.9978	.9983	.9987	.9991	.9996	0 1 1	2 2 3	3 3 4

Antilogarithms—Common System

L	A										B		
	0	1	2	3	4	5	6	7	8	9	1 2 3	4 5 6	7 8 9
.00	1000	1002	1005	1007	1009	1012	1014	1016	1019	1021	0 0 1	1 1 1	2 2 2
.01	1023	1026	1028	1030	1033	1035	1038	1040	1042	1045	0 0 1	1 1 1	2 2 2
.02	1047	1050	1052	1054	1057	1059	1062	1064	1067	1069	0 0 1	1 1 1	2 2 2
.03	1072	1074	1076	1079	1081	1084	1086	1089	1091	1094	0 0 1	1 1 1	2 2 2
.04	1096	1099	1102	1104	1107	1109	1112	1114	1117	1119	0 1 1	1 1 2	2 2 2
.05	1122	1125	1127	1130	1132	1135	1138	1140	1143	1146	0 1 1	1 1 2	2 2 2
.06	1148	1151	1153	1156	1159	1161	1164	1167	1169	1172	0 1 1	1 1 2	2 2 2
.07	1175	1178	1180	1183	1186	1189	1191	1194	1197	1199	0 1 1	1 1 2	2 2 2
.08	1202	1205	1208	1211	1213	1216	1219	1222	1225	1227	0 1 1	1 1 2	2 2 3
.09	1230	1233	1236	1239	1242	1245	1247	1250	1253	1256	0 1 1	1 1 2	2 2 3
.10	1259	1262	1265	1268	1271	1274	1276	1279	1282	1285	0 1 1	1 1 2	2 2 3
.11	1288	1291	1294	1297	1300	1303	1306	1309	1312	1315	0 1 1	1 2 2	2 2 3
.12	1318	1321	1324	1327	1330	1334	1337	1340	1343	1346	0 1 1	1 2 2	2 2 3
.13	1349	1352	1355	1358	1361	1365	1368	1371	1374	1377	0 1 1	1 2 2	2 3 3
.14	1380	1384	1387	1390	1393	1396	1400	1403	1406	1409	0 1 1	1 2 2	2 3 3
.15	1413	1416	1419	1422	1426	1429	1432	1435	1439	1442	0 1 1	1 2 2	2 3 3
.16	1445	1449	1452	1455	1459	1462	1466	1469	1472	1476	0 1 1	1 2 2	2 3 3
.17	1479	1483	1486	1489	1493	1496	1500	1503	1507	1510	0 1 1	1 2 2	2 3 3
.18	1514	1517	1521	1524	1528	1531	1535	1538	1542	1545	0 1 1	1 2 2	2 3 3
.19	1549	1552	1556	1560	1563	1567	1570	1574	1578	1581	0 1 1	1 2 2	3 3 3
.20	1585	1589	1592	1596	1600	1603	1607	1611	1614	1618	0 1 1	1 2 2	3 3 3
.21	1622	1626	1629	1633	1637	1641	1644	1648	1652	1656	0 1 1	2 2 2	3 3 3
.22	1660	1663	1667	1671	1675	1679	1683	1687	1690	1694	0 1 1	2 2 2	3 3 3
.23	1698	1702	1706	1710	1714	1718	1722	1726	1730	1734	0 1 1	2 2 2	3 3 4
.24	1738	1742	1746	1750	1754	1758	1762	1766	1770	1774	0 1 1	2 2 2	3 3 4
.25	1778	1782	1786	1791	1795	1799	1803	1807	1811	1816	0 1 1	2 2 2	3 3 4
.26	1820	1824	1828	1832	1837	1841	1845	1849	1854	1858	0 1 1	2 2 3	3 3 4
.27	1862	1866	1871	1875	1879	1884	1888	1892	1897	1901	0 1 1	2 2 3	3 3 4
.28	1905	1910	1914	1919	1923	1928	1932	1936	1941	1945	0 1 1	2 2 3	3 4 4
.29	1950	1954	1959	1963	1968	1972	1977	1982	1986	1991	0 1 1	2 2 3	3 4 4
.30	1995	2000	2004	2009	2014	2018	2023	2028	2032	2037	0 1 1	2 2 3	3 4 4
.31	2042	2046	2051	2056	2061	2065	2070	2075	2080	2084	0 1 1	2 2 3	3 4 4
.32	2089	2094	2099	2104	2109	2113	2118	2123	2128	2133	0 1 1	2 2 3	3 4 4
.33	2138	2143	2148	2153	2158	2163	2168	2173	2178	2183	0 1 1	2 2 3	3 4 4
.34	2188	2193	2198	2203	2208	2213	2218	2223	2228	2234	1 1 2	2 3 3	4 4 5
.35	2239	2244	2249	2254	2259	2265	2270	2275	2280	2286	1 1 2	2 3 3	4 4 5
.36	2291	2296	2301	2307	2312	2317	2323	2328	2333	2339	1 1 2	2 3 3	4 4 5
.37	2344	2350	2355	2360	2366	2371	2377	2382	2388	2393	1 1 2	2 3 3	4 4 5
.38	2399	2404	2410	2415	2421	2427	2432	2438	2443	2449	1 1 2	2 3 3	4 4 5
.39	2455	2460	2466	2472	2477	2483	2489	2495	2500	2506	1 1 2	2 3 3	4 5 5
.40	2512	2518	2523	2529	2535	2541	2547	2553	2559	2564	1 1 2	2 3 4	4 5 5
.41	2570	2576	2582	2588	2594	2600	2606	2612	2618	2624	1 1 2	2 3 4	4 5 5
.42	2630	2636	2642	2649	2655	2661	2667	2673	2679	2685	1 1 2	2 3 4	4 5 6
.43	2692	2698	2704	2710	2716	2723	2729	2735	2742	2748	1 1 2	3 3 4	4 5 6
.44	2754	2761	2767	2773	2780	2786	2793	2799	2805	2812	1 1 2	3 3 4	4 5 6
.45	2818	2825	2831	2838	2844	2851	2858	2864	2871	2877	1 1 2	3 3 4	5 5 6
.46	2884	2891	2897	2904	2911	2917	2924	2931	2938	2944	1 1 2	3 3 4	5 5 6
.47	2951	2958	2965	2972	2979	2985	2992	2999	3006	3013	1 1 2	3 3 4	5 5 6
.48	3020	3027	3034	3041	3048	3055	3062	3069	3076	3083	1 1 2	3 4 4	5 6 6
.49	3090	3097	3105	3112	3119	3126	3133	3141	3148	3155	1 1 2	3 4 4	5 6 6

Antilogarithms—Common System (Contd.)

L	A										B		
	0	1	2	3	4	5	6	7	8	9	1 2 3	4 5 6	7 8 9
.50	3162	3170	3177	3184	3192	3199	3206	3214	3221	3228	1 1 2	3 4 4	5 6 7
.51	3236	3243	3251	3258	3266	3273	3281	3289	3296	3304	1 2 2	3 4 5	5 6 7
.52	3311	3319	3327	3334	3342	3350	3357	3365	3373	3381	1 2 2	3 4 5	5 6 7
.53	3388	3396	3404	3412	3420	3428	3436	3443	3451	3459	1 2 2	3 4 5	6 6 7
.54	3467	3475	3483	3491	3499	3508	3516	3524	3532	3540	1 2 2	3 4 5	6 6 7
.55	3548	3556	3565	3573	3581	3589	3597	3606	3614	3622	1 2 2	3 4 5	6 7 7
.56	3631	3639	3648	3656	3664	3673	3681	3690	3698	3707	1 2 3	3 4 5	6 7 8
.57	3715	3724	3733	3741	3750	3758	3767	3776	3784	3793	1 2 3	3 4 5	6 7 8
.58	3802	3811	3819	3828	3837	3846	3855	3864	3873	3882	1 2 3	4 4 5	6 7 8
.59	3890	3899	3908	3917	3926	3936	3945	3954	3963	3972	1 2 3	4 5 5	6 7 8
.60	3981	3990	3999	4009	4018	4027	4036	4046	4055	4064	1 2 3	4 5 6	6 7 8
.61	4074	4083	4093	4102	4111	4121	4130	4140	4150	4159	1 2 3	4 5 6	7 8 9
.62	4169	4178	4188	4198	4207	4217	4227	4236	4246	4256	1 2 3	4 5 6	7 8 9
.63	4266	4276	4285	4295	4305	4315	4325	4335	4345	4355	1 2 3	4 5 6	7 8 9
.64	4365	4375	4385	4395	4406	4416	4426	4436	4446	4457	1 2 3	4 5 6	7 8 9
.65	4467	4477	4487	4498	4508	4519	4529	4539	4550	4560	1 2 3	4 5 6	7 8 9
.66	4571	4581	4592	4603	4613	4624	4634	4645	4656	4667	1 2 3	4 5 6	7 9 10
.67	4677	4688	4699	4710	4721	4732	4742	4753	4764	4775	1 2 3	4 5 7	8 9 10
.68	4786	4797	4808	4819	4831	4842	4853	4864	4875	4887	1 2 3	4 6 7	8 9 10
.69	4898	4909	4920	4932	4943	4955	4966	4977	4989	5000	1 2 3	5 6 7	8 9 10
.70	5012	5023	5035	5047	5058	5070	5082	5093	5105	5117	1 2 4	5 6 7	8 9 11
.71	5129	5140	5152	5164	5176	5188	5200	5212	5224	5236	1 2 4	5 6 7	8 10 11
.72	5248	5260	5272	5284	5297	5309	5321	5333	5346	5358	1 2 4	5 6 7	9 10 11
.73	5370	5383	5395	5408	5420	5433	5445	5458	5470	5483	1 3 4	5 6 8	9 10 11
.74	5495	5508	5521	5534	5546	5559	5572	5585	5598	5610	1 3 4	5 6 8	9 10 12
.75	5623	5636	5649	5662	5675	5689	5702	5715	5728	5741	1 3 4	5 7 8	9 10 12
.76	5754	5768	5781	5794	5808	5821	5834	5848	5861	5875	1 3 4	5 7 8	9 11 12
.77	5888	5902	5916	5929	5943	5957	5970	5984	5998	6012	1 3 4	5 7 8	10 11 12
.78	6026	6039	6053	6067	6081	6095	6109	6124	6138	6152	1 3 4	6 7 8	10 11 13
.79	6166	6180	6194	6209	6223	6237	6252	6266	6281	6295	1 3 4	6 7 9	10 11 13
.80	6310	6324	6339	6353	6368	6383	6397	6412	6427	6442	1 3 4	6 7 9	10 12 13
.81	6457	6471	6486	6501	6516	6531	6546	6561	6577	6592	2 3 5	6 8 9	11 12 14
.82	6607	6622	6637	6653	6668	6683	6699	6714	6730	6745	2 3 5	6 8 9	11 12 14
.83	6761	6776	6792	6808	6823	6839	6855	6871	6887	6902	2 3 5	6 8 9	11 13 14
.84	6918	6934	6950	6966	6982	6998	7015	7031	7047	7063	2 3 5	6 8 10	11 13 15
.85	7079	7096	7112	7129	7145	7161	7178	7194	7211	7228	2 3 5	7 8 10	12 13 15
.86	7244	7261	7278	7295	7311	7328	7345	7362	7379	7396	2 3 5	7 8 10	12 13 15
.87	7413	7430	7447	7464	7482	7499	7516	7534	7551	7568	2 3 5	7 9 10	12 14 16
.88	7586	7603	7621	7638	7656	7674	7691	7709	7727	7745	2 4 5	7 9 11	12 14 16
.89	7762	7780	7798	7816	7834	7852	7870	7889	7907	7925	2 4 5	7 9 11	13 14 16
.90	7943	7962	7980	7998	8017	8035	8054	8072	8091	8110	2 4 6	7 9 11	13 15 17
.91	8128	8147	8166	8185	8204	8222	8241	8260	8279	8299	2 4 6	8 9 11	13 15 17
.92	8318	8337	8356	8375	8395	8414	8433	8453	8472	8492	2 4 6	8 10 12	14 15 17
.93	8511	8531	8551	8570	8590	8610	8630	8650	8670	8690	2 4 6	8 10 12	14 16 18
.94	8710	8730	8750	8770	8790	8810	8831	8851	8872	8892	2 4 6	8 10 12	14 16 18
.95	8913	8933	8954	8974	8995	9016	9036	9057	9078	9099	2 4 6	8 10 12	15 17 19
.96	9120	9141	9162	9183	9204	9226	9247	9268	9290	9311	2 4 6	8 11 13	15 17 19
.97	9333	9354	9376	9397	9419	9441	9462	9484	9506	9528	2 4 7	9 11 13	15 17 20
.98	9550	9572	9594	9616	9638	9661	9683	9705	9727	9750	2 4 7	9 11 13	16 18 20
.99	9772	9795	9817	9840	9863	9886	9908	9931	9954	9977	2 5 7	9 11 14	16 18 20

APPENDIX E*

PART 1—SHORT METHOD OF COMPUTING THE ARITHMETIC AVERAGE

The "Assumed Average"

This method of shortening the process of computing the arithmetic average is based on the fact that the sum of the variations of the several items in the group from the mean, taken algebraically, is equal to zero. It consists of the following steps:

1. Select an "assumed" average.
2. Take the deviations of the several items in the group from the "assumed" average.
3. Summate these deviations algebraically—*i.e.*, giving recognition to positive and negative numbers.
4. Divide this sum by the number of items, also stated as the sum of the frequencies or the sum of the weights. If the sum is zero, the "assumed" average is the correct or true average and no further computation need be made.
5. Add the quotient obtained in No. 4 algebraically to the assumed average. That is, if it is a negative number, it should be subtracted from the assumed average since it is obvious in this case that the assumed average is too large. On the other hand, if it is positive, it should be added to the assumed average.

Table 1 below will clarify the process.

In this illustration, the result of the first step indicated above is noted in column "c"; the second step in column "d"; and the third step at the foot of column "d." The fourth and fifth steps are indicated in the caption.

Avoiding the Algebraic Feature of the Assumed Average.—To avoid the algebraic feature of the process the lowest or highest item in the group may be taken as the assumed average in which case the negative quantities will be eliminated. If the smallest figure is taken, we subtract it from the several items in the group to get the deviations, and the final result is added to the assumed average. If the largest figure is taken, we subtract the several items in the group from it to get the deviations, and the final result is subtracted from the assumed average.

Tables 2 and 3 below are illustrative.

* The separate parts of this Appendix are independent of each other, the development of each one bearing no relation to the others.

AVERAGE WEIGHT OF A BALE OF COTTON

Year	Weight	Assumed Average	Deviations from Assumed Average
a	b	c	d
1910	475	482	− 7
1911	480	482	− 2
1912	482	482	0
1913	486	482	+ 4
1914	484	482	+ 2
		Algebraic Sum	− 3

TABLE 1.—The Use of an "Assumed" Average in Computing the Arithmetic Average

Dividing (−3) by 5, the number of years, we get (−0.6) as the average deviation from the assumed average. Adding this algebraically to the assumed average we get a result of 481.4, the correct or true average, i.e., 482 + (−0.6) = 481.4.

(Data same as in Table 1)

Year	Weight	Assumed Average	Deviations
a	b	c	d
1910	475	475	0
1911	480	475	5
1912	482	475	7
1913	486	475	11
1914	484	475	9
		Sum	32

TABLE 2.—Taking the Smallest Number in the Series as the "Assumed" Average.

Dividing 32 by 5, the number of years, we get 6.4 as the average deviation from the assumed average. Adding this to the assumed average the result is 481.4, the true average.

THE "ASSUMED" AVERAGE AND FREQUENCY SERIES.—This method may be applied to frequency series where the items are grouped in classes as illustrated in Table 4 below.

We might also take the largest wage group or any other figure as the assumed average.

THE STEP METHOD.—A modification of the "assumed average" method in frequency series makes use of the variations by classes rather than by values of the mid-points of the class-groups. This is called the "step"

(Data same as in Table 1)

Year	Weight	Assumed Average	Deviations
a	b	c	d
1910	475	486	11
1911	480	486	6
1912	482	486	4
1913	486	486	0
1914	484	486	2
		Sum	23

TABLE 3.—Taking the Largest Item in the Series as the "Assumed" Average.

Dividing 23 by 5, the number of years, we get 4.6 as the average deviation from the assumed average. Subtracting this from the assumed average (486–4.6=481.4) the result is 481.4, the true average.

WAGE STATISTICS

BASE RATES OF PAY—WEEKLY BASIS
(To Nearest Dollar)

Wage Groups	No. of Employes	Mid-Points of Groups	Assumed Average	Deviations	Product (e × b)
a	b	c	d	e	f
$10 under $15	10	$12.00	$12.00	0	$ 0.00
15 " 20	15	17.00	12.00	5	75.00
20 " 25	50	22.00	12.00	10	500.00
25 " 30	12	27.00	12.00	15	180.00
30 " 35	8	32.00	12.00	20	160.00
35 and over	5	37.00	12.00	25	125.00
Total	100				$1,040.00

TABLE 4.—Illustrating the Use of the "Assumed" Average in Computing the Arithmetic Average in a Frequency Series.

Dividing $1,040 by 100, we get $10.40 as the average deviation from the assumed average. Hence, $12.00, the assumed average, plus $10.40, equals $22.40, the correct or true average.

method, each class-group being considered a "step." Table 5 below will illustrate.

This process differs from the other only in the fact that in this case we first find the step-deviation and then by multiplying this by the class-interval we convert it to terms of the unit of measurement.

SUMMARY.—This method has no relation to the character of the average as such. It is merely an arithmetical device used to abbreviate the arithmetic operation involved in computing the arithmetic average and has no effect whatever on the final results obtained.

WAGE STATISTICS

BASE RATES OF PAY—WEEKLY BASIS

Wage Groups	No. of Employes	Assumed Average	Step-Deviations	Products (b × d)
a	b	c	d	e
$10 under $15	10	$12.00	0	$ 0
15 " 20	15	12.00	1	15
20 " 25	50	12.00	2	100
25 " 30	12	12.00	3	36
30 " 35	8	12.00	4	32
35 and over	5	12.00	5	25
Total	100			$208

TABLE 5.—The "Step" Method of Computing the Arithmetic Average.

Dividing 208 by 100 we get 2.08, the average step-deviation. But each step is $5.00 (class interval) in width and consequently 2.08 steps is equivalent to ($5 × 2.08), or $10.40, the average deviation in dollars and cents. Hence, the true average is $12.00 plus $10.40, or $22.40.

In the illustrations used above, the assumed average has been an actual item of the series which, of course, need not be the case. We take any number as the assumed average which seems preferable in the specific case whether it appears in the group or not. We aim to select the specific figure which will make the process of computation easiest.

Since in elementary business statistics, statistical computations are done with computing machines it is rarely that we find occasion to use this short-cut method.

PART 2—THE PEARSONIAN COEFFICIENT OF CORRELATION

The most exact measure of correlation is found in the computation of a coefficient of correlation. The method of computation most generally used is that worked out by Karl Pearson, the English biological statistician. The following formula is used in the computation of this coefficient:

$$r = \frac{\Sigma (xy)}{n\,(S.D._x)\,(S.D._y)}$$

in which, r = coefficient of correlation.
x = deviations of the items of one series from the arithmetic average.
y = deviations of the items of the other series from the arithmetic average.
$(S.D._x)$ = standard deviation of the one series.
$(S.D._y)$ = standard deviation of the other series.
n = number of pairs of items.

Department	Total Payroll (hundreds of dollars)	Deviations from Arithmetic Average (x)	Deviations Squared (x^2)	(x y)
a	b	c	d	e
Accounting	2	− 25	625	1,550
Engineering	23	− 4	16	72
Purchasing	24	− 3	9	18
Sales	36	+ 9	81	279
Production	50	+ 23	529	1,242
Total	135	.....	1,260	3,161
Average	27			

$$\text{S. D.}_x = \sqrt{\frac{\Sigma\,(\text{deviations})^2}{n}} = \sqrt{\frac{1260}{5}} = \sqrt{252} = 16 -$$

Department	Total Period of Service (hundreds of months)	Deviations from Arithmetic Average (y)	Deviations Squared (y^2)
a	b	c	d
Accounting	3	− 62	3,844
Engineering	47	− 18	324
Purchasing	59	− 6	36
Sales	96	+ 31	961
Production	119	+ 54	2,916
Total	324	.....	8,081
Average	65		

$$\text{S. D.}_y = \sqrt{\frac{8081}{5}} = \sqrt{1616} = 40 +$$

$$r = \frac{\Sigma xy}{n\,(\text{S.D.}_x)\,(\text{S.D.}_y)} = \frac{3161}{5 \times 16 \times 40} = 0.9 +$$

TABLE 1.—The Sum-Product Method of Computing the Coefficient of Correlation in Ungrouped Series.

WAGES IN CENTS PER HOUR (Y-Series)	PERIOD OF SERVICE IN YEARS (X-Series)									Σ x′ y′
	Under 1 Year	1 – 2	2 – 3	3 – 4	4 – 5	5 – 6	6 – 7	7 – 8	8 – 9	
a	b	c	d	e	f	g	h	i	j	k
35 – 39	8 27 216	4 13 52	3	−4 1 −4						264
40 – 44	6 34 204	3 19 57	10							261
45 – 49	4 25 100	2 49 98	20	−2 1 −2	−4 1 −4					192
50 – 54	2 10 20	1 52 52	22	−1 10 −10	−2 6 −12	−3 2 −6				44
55 – 59	9	53	64	14	35	25	8	2	1	
60 – 64	−2 11 −22	−1 28 −28	8	1 7 7	2 22 44	3 29 87	4 14 56	5 7 35	6 2 12	191
65 – 69	−4 4 −16	−2 8 −16	17	2 3 6	4 2 8	6 9 54	8 9 72	10 3 30	12 1 12	150
70 – 74		−3 9 −27	3	3 3 9	6 5 30	9 2 18	12 3 36	15 1 15	18 2 36	117
75 – 79	−8 1 −8	−4 4 −16			8 7 56	12 1 12		20 1 20	24 2 48	112

TABLE 2.—Correlation Table Illustrating the Sum-Product Method of Computing the Coefficient of Correlation.

It will be noted that all of the items of the right-hand member of this equation are familiar. Table 1 above together with the accompanying text will clarify the process in the ungrouped data and Tables 2, 3 and 4 in the grouped series. It will be noted that in the latter case the process is shortened materially by the use of the assumed average and class deviations.

In this method of determining correlation, a coefficient of + 1 indicates perfect positive correlation, —1, perfect negative correlation, and zero the absence of correlation altogether. The point at which the correlation becomes significant will vary with the character of the problem and the difficulty of determining this fact is perhaps the greatest obstacle in the practical application of this method of statistical analysis. This point is discussed at some length in the references listed on page 470.

As a matter of practice, however, the coefficient of correlation will not be used extensively in elementary business statistics. The use of the straight line curve while less precise is sufficiently exact for practical purposes and will generally serve our needs in the more elementary work.

Y-Series					X-Series					
Wage Groups (m)	No. of Employes (f)	Step Deviations (d'_y)	($f\,d'_y$)	$f\,(d'_y)^2$	Service Group (m)	No. of Employes (f)	Step Deviations (d'_x)	($f\,d'_x$)	$f\,(d'_x)^2$	($x'\,y'$)
a	b	c	d	e	f	g	h	i	j	k
37	44	−4	−176	704	0.5	121	−2	−242	484	264
42	63	−3	−189	567	1.5	235	−1	−235	235	261
47	96	−2	−192	384	2.5	147	0	0	0	192
52	102	−1	−102	102	3.5	39	+1	+ 39	39	44
57	211	0	0	0	4.5	78	+2	+156	312	
62	128	+1	+128	128	5.5	68	+3	+204	612	191
67	56	+2	+112	224	6.5	34	+4	+136	544	150
72	28	+3	+ 84	252	7.5	14	+5	+ 70	350	117
77	16	+4	+ 64	256	8.5	8	+6	+ 48	288	112
Total	744 (n)		−271	2,617	Total	744 (n)		+176	2,864	1,331

TABLE 3.—The Sum-Product Method of Computing the Coefficient of Correlation in a Grouped Series.

To shorten the process, the assumed average or arbitrary origin and the step or class deviations have been used. In such cases d′ instead of d is used to designate the deviations.

Let c = difference between the assumed and correct average.
s = S.D. from the assumed average.

Then $(S.D.) = \sqrt{s^2 - c^2}$

$$\text{Now } c_x = \frac{\Sigma fd'_x}{n} = \frac{176}{744} = .237$$

$$c_y = \frac{-271}{744} = -.364$$

$$s_x^2 = \frac{\Sigma fd'^2_x}{n} = \frac{2864}{744} = 3.85$$

$$s_y^2 = \frac{\Sigma fd'^2_y}{n} = \frac{2617}{744} = 3.52$$

$$\text{Hence } (S.D._x) = \sqrt{s_x^2 - c_x^2}$$
$$= \sqrt{3.85 - (.237)^2}$$
$$= \sqrt{3.85 - .0562}$$
$$= \sqrt{3.7938}$$
$$= 1.96$$

$$\text{And } (S.D._y) = \sqrt{s_y^2 - c_y^2}$$
$$= \sqrt{3.52 - (-.364)^2}$$
$$= \sqrt{3.52 - .1325}$$
$$= \sqrt{3.3875}$$
$$= 1.84$$

When the assumed average or arbitrary origin and the step or class deviations are used the formula becomes,

$$r = \frac{\frac{\Sigma x'y'}{n} - c_x c_y}{(S.D._x)\,(S.D._y)}.$$

Substituting in this equation,

$$r = \frac{\frac{1331}{744} - (.237)(-.364)}{(1.96)\,(1.84)}$$

$$= \frac{1.7889 + .0863}{3.5164}$$

$$= \frac{1.8752}{3.5164}$$

$$= +.5$$

Classes in Y-Series			Distribution by X-Classes			$\Sigma x'y'$
	37	f	27	13	1	
		d'_x	2	1	−1	
		d'_y	4	4	4	
		$d'_x d'_y$	8	4	−4	
		$fd'_x d'_y$	216	52	−4	264
	42	f	34	19		
		d'_x	2	1		
		d'_y	3	3		
		$d'_x d'_y$	6	3		
		$fd'_x d'_y$	204	57		261

TABLE 4.—Illustrating the Computation of the x'y' Items Given in Column "K" of Table 3.

There are 27 employes in the 37-cent wage group with a period of service of less than one year. By an inspection of Table 2 it will be seen that the class deviation of this item from the assumed average of the X-series (*i.e.*, 2.5) is 2, and from the assumed average of the Y-series (*i.e.*, 57), 4. Hence, the continued product of 27, 2, and 4 ($27 \times 2 \times 4 = 216$) will give the total deviations of these 27 units. The deviations of the other groups are determined in similar manner. The $d'_x d'_y$ products appear at the top of the boxes in Table 2 and the $fd'_x d'_y$ products at the bottom.

PART 3—SECULAR TREND—METHOD OF LEAST SQUARES

The method of least squares was referred to on page 478 as a third method of determining the secular trend. This method consists of fitting the line of least squares to the actual data, so-called because the sum of the squares of the vertical deviations from this line is less than the sum of the squares of like deviations from any other line. The method of determining this line consists of the following steps which are illustrated in Tables 1 and 2:

1. Compute the average of the actual data. This is 3,485 divided by 5, or 697 for the data of Table 1. This is the mid-point of the line of trend. For convenience of computation the yearly instead of the monthly data are used.
2. Find the mid-point in time for the series for which the trend is to be determined. In Table 1 this is the middle of the year 1921, or the point between June and July of that year.
3. Determine the annual increment as follows: Take the plus and minus deviations in years from the mid-point in time (column "c" of Table 1) and square them (column "e" of Table 1). Also take the product of these deviations and the actual data (column "d" of Table 1). Finally, divide the sum of the products by the sum of the squares of the deviations. In the form of a formula, $\frac{\Sigma xy}{\Sigma x^2}$ = annual increment. Substituting from Table 1, we have $\frac{1637}{10} = 163.7$ as annual increment.

One-twelfth of this, or 13.7 is the monthly increment.

4. Determine the ordinates of secular trend. This is accomplished by subtracting the annual or monthly increments successively from the average to obtain the ordinates for periods earlier than the mid-point and adding them to obtain ordinates for periods later than the mid-point. In the data of Table 1, since the mid-point lies between June and July for the year 1921, we must subtract one-half of the monthly increment, 13.7, from the average, 697, to get the ordinate for June, 690.2. Then by subtracting the monthly increment successively from this figure for June, we obtain the ordinates for the months from June 1921 back to January 1919; and by adding the monthly increment we obtain the ordinates for the months from June 1921 to December 1923. Thus 690.2 minus 13.7 gives 676.5, the ordinate for May 1921; and 690.2 plus 13.7 gives 703.9 the ordinate for August 1921. The monthly ordinates are shown in Table 2. Of course, only two points are required to construct the line of trend but the monthly ordinates will be required in eliminating the influence of the trend. The line of least squares for monthly stocks of gasoline is plotted in Figure 115, page 477.

GASOLINE STOCKS

Millons of Gallons

Years	Actual Data Monthly Average (y)	Deviation from Mid-point (x)	Product of Deviation and Actual Data (x y)	Deviations Squared (x^2)
a	b	c	d	e
1919	472	− 2	− 944	4
1920	464	− 1	− 946	1
1921	631	0	0	0
1922	791	+ 1	791	1
1923	1,127	+ 2	2,254	4
Total	3,485	0	1,637	10

TABLE 1.—Computations in Determining Secular Trend by Method of Least Squares

GASOLINE STOCKS

Millions of Gallons

Months	1919	1920	1921	1922	1923
a	b	c	d	e	f
January	292.9	457.3	621.7	786.1	950.5
February	306.6	471.0	635.4	799.8	964.2
March	320.3	484.7	649.1	813.5	977.9
April	334.0	498.4	662.8	827.2	991.6
May	347.7	512.1	676.5	840.9	1,005.3
June	361.4	525.8	690.2	854.6	1,019.0
July	375.1	539.5	703.9	868.3	1,032.7
August	388.8	553.2	717.6	882.0	1,046.4
September	402.5	566.9	731.3	895.7	1,060.1
October	416.2	580.6	745.0	909.4	1,073.8
November	429.9	594.3	758.7	923.1	1,087.5
December	443.6	608.0	772.4	936.8	1,101.2

TABLE 2.—Ordinates of Secular Trend by Method of Least Squares

PART 4—SEASONAL INDEX—METHOD OF LINK RELATIVES

This method is somewhat more involved in process and more burdensome in computations than the two methods developed on pages 478 and 479. It consists of the following steps:

1. Link relatives are computed from the original data. That is, the item for each month is expressed as a percentage of the item for the month immediately preceding. Thus, 458, the item for February 1919, is expressed as a percentage of 383, the item for January, to get the figure 119.6, the link relative for February. March is expressed as a percentage of February; April, of March, and so on throughout the period. (Columns "b" to "f" of Table 1.)

2. The median link relative is located for each month. This is accomplished by arraying the items for a given month in order of magnitude and locating the median value in the usual manner. (Column "g" in Table 1 below.)

3. Chain relatives are computed from the median link relatives. This step is designed to convert the relatives to a common base. Usually January is taken as the common base, the chain relative of which would be 100, and the process consists then of multiplying the link relative of a given month by the chain relative of the month preceding. February link relative is multiplied by the chain relative for January; the March link relative is multiplied by the chain relative for February; and so on for the other months. (Column "h" in Table 1 below.)

4. The chain relatives are adjusted to eliminate the influence of the secular trend. The effect of the secular trend is represented by the difference between the chain for January (100) at the beginning of the series and that determined by multiplying the December chain relative by the January link relative which in Table 1 is 130.5. This difference represents the compounding of an error over the 12 months' which—letting d represent the monthly error —is eliminated by dividing the February chain relative by $(1 + d)$, March by $(1 + d)^2$, and so on. These computations are readily made by means of logarithms.[1] If the difference to be adjusted is small, we may divide 30.5 by 12 and subtract the result from February, twice the result from March, and so forth, with but slight error.[2] Column "i" of Table 1 gives the corrected chain relatives.

5. The final step is to shift the base from January to the average for the year. This is accomplished by dividing each item in column "i" by the average of the items, in column "i." The final index is listed in column "j" of Table 1.

GASOLINE STOCKS—MILLIONS OF GALLONS

Months	1919	1920	1921	1922	1923	Median Link Relatives	Chain Relatives	Corrected Chain Relatives	Seasonal Index
a	b	c	d	e	f	g	h	i	j
January	129.0	115.4	123.8	120.5	113.5	120.5	100.0	100.0	101.8
February	119.6	109.1	119.1	114.3	112.7	114.3	114.3	111.8	113.9
March	119.2	111.2	104.7	105.8	111.4	111.2	127.1	121.6	123.8
April	108.7	102.9	104.8	104.6	106.1	104.8	133.2	124.7	126.9
May	100.0	89.8	107.2	95.9	99.5	99.5	132.5	121.3	123.4
June	100.0	87.2	93.8	96.3	95.1	95.1	126.0	112.8	114.9
July	86.7	81.9	91.1	93.7	92.2	91.1	114.8	100.6	102.4
August	84.5	78.2	83.0	91.1	90.5	84.5	97.0	83.1	84.6
September	85.3	89.2	90.7	98.0	92.3	90.7	87.9	73.7	75.1
October	95.4	104.5	88.5	104.9	97.3	97.3	85.5	70.1	71.4
November	106.8	117.9	108.8	107.3	104.0	107.3	91.7	73.5	74.8
December	118.3	130.1	118.1	113.8	109.1	118.1	108.3	84.9	86.5
							Jan. 130.5	Av. 98.2	

TABLE 1.—Computation of Seasonal Variation by Method of Link Relatives

PART 5—CHECKING STATISTICAL DATA

Inaccurate figures in statistical work are worse than none at all unless the presence of the errors is recognized and due allowance made for them. The statistician whose work is inaccurate has no more excuse for existence than the dentist who on pulling the wrong tooth, apologetically exclaims, "Never mind, I'll get the right one the next time." Even small errors are detrimental. They may not seriously affect the results but may throw dis-

[1] See F. C. Mills' "Statistical Methods," pages 319-321.
[2] See Harry Jerome "Statistical Method," page 242.

credit on the work as a whole. Consequently, it is highly important that checks be used liberally in every phase of statistical work.

The method used to check statistical data will vary with the character of the data and the facts surrounding their development. The following are some of the methods most commonly used in practical statistical work in a business organization:

1. *Repetition*—This method of checking applies particularly to computations and consists primarily of performing the operation more than once independently and comparing the results of the separate operations. We may, for example, compute average rates of pay or average prices, and then have the process repeated independently by the same person or preferably by a different person. The repetition may produce a detailed check in which case the second operation is designed to produce as accurate results as the first, such as for example, when the operations of the Comptometer are checked by having a second operator perform the same identical operations. Or, we may get a rougher check as when the operations are repeated by means of the slide rule.

It is evident that the efficacy of this method depends on the probability that exactly the same error will not be made in two independent operations. This affords a very good check—one that is quite satisfactory for practical purposes.

2. *Reversing the Operations*—This method consists merely in taking the results obtained from the operation and working back to get the known quantities on which the operation was based, as, for example, when a quotient is multiplied by the divisor and the result checked with the dividend. It is confined for the most part to certain types of cases such as the checking of arithmetic averages, percentages, ratios, and the simple processes of subtraction, multiplication, and division.

3. *Checking by Totals*—This method consists merely in checking totals with one another and making no specific check on the details. Thus we may so arrange the data in tabular form as to produce cross-totals as shown in Table 1, on page 248. Or, we may have totals on separate statements which must be equal. Or, again, we may have a series of sub-totals, which combined must equal a main total as shown in the same table. In this latter case sometimes the sub-totals because of less exactness, such as figuring only to the nearest dollar, or 100, or 1,000, will not add up to the main totals. This is particularly true with percentages. In such cases, their exact totals should be recorded and the disparity explained in a footnote.

Obviously this is not a guarantee of accuracy of detail, as compensating errors may not be reflected in the total.

4. *Spot-Check*—This method is in the nature of sample checking. Certain items only are selected and checked and the accuracy of the group is judged from these. When we are making hundreds of computations, for example, we may select 10 or 20 at random from each hundred and care-

fully check these, thus getting a rough estimate of the accuracy of the entire group.

It is evident that this method is closely associated with statistical sampling and that its efficacy depends on the representativeness of the *samples* or *"spots"* checked. It is of course not a 100% check but it is sufficiently accurate for many purposes in practical work.

5. *Reading Back*—This method consists of reading back the numbers from one statement or record and comparing them with those on another statement or record with which they are supposed to agree. In using it, one person—preferably the more responsible one—checks the copy while another person reads from the original data.

6. *Similar Data*—This method consists simply of comparing data with corresponding items for the previous month or year or other time period or with similar data on independent statements. Thus we may check expenses, sales, labor turnover, and so forth, for one month with corresponding items for previous months this year or last. Or, we may check facts of our own company against similar facts for other companies or for the community.

This method should be invariably applied to important figures before they are released, particularly those which represent conclusions.

7. *Indirection*—This method is likewise based on the relationship between the data being checked and certain other data, but in this case they are checked with data to which they bear a somewhat fixed relationship rather than with data with which they strictly correspond. Thus there is a definite relationship between sales deliveries and sales orders, between labor costs and output, between the payroll and the number of employes, and so forth, and rough checks may be made on the basis of these relationships. This method, also, is only a broad check and is serviceable primarily as a prevention against serious error.

8. *Common Knowledge*—This method is similar to the last two except that it depends on general familiarity with the kind of data being checked. We know generally what the average rate of pay, the number of employes, the selling expense, the labor costs, and so forth, should be and figures widely different will be readily detected by inspection. It is by this method of checking that the experienced executive can quickly and with almost uncanny incisiveness detect errors in statistical statements which come to him.

This method is extremely serviceable to the experienced statistician. There are many cases in practical work where we catch errors simply because the figures "don't look right."

9. *Compensating and Cumulative Errors*—It is also serviceable to ascertain, if possible, whether errors which cannot be eliminated tend to compensate—*i.e.*, offset each other—or tend to cumulate. If the former is the case, they are not so serious and may in some instances be disregarded, while if errors cumulate they may invalidate the data altogether.

PART 6—QUESTIONNAIRES

(Example A)

Common Uses of English for Communication

Interviews

1. **With customers or patrons, if you are a business or professional man, or a representative of a service corporation.**

   Do you use English in this way?.....Frequently?.....Do you find that this use of English presents any outstanding difficulty?.....

   Write in the space below any details you care to give regarding difficulties you may experience. (*Many have trouble in finding out quickly just what the customer or patron wants; in giving necessary explanations without waste of time; in refusing undue concessions or making collections without losing the customer's good will, etc.*)

   .....

   .....

2. **When making oral reports to an official superior.**

   Do you use English in this way?.....Frequently?.....Do you find that this use of English presents any outstanding difficulty?.....

   Write in the space below any details you care to give regarding difficulties you may experience. (*Difficulty is sometimes found in making the story definite and complete without too much detail, arranging it so that the main points stand out clearly; speaking in a manner that is business like and easy, etc.*)

   .....

   .....

3. **Giving oral instructions to subordinates or agents.**

   Do you use English in this way?.....Frequently?.....Do you find that this use of English presents any outstanding difficulty?.....

   Write in the space below any details you care to give regarding difficulties you may experience. (*To do this well involves: using words that are easily understood by the person addressed; exact statements; brevity; a manner that is energetic without being impatient or curt, etc.*)

   .....

   .....

Conversation

1. **With strangers or casual acquaintances.**

   Do you use English in this way?.....Frequently?.....Do you find that this use of English presents any outstanding difficulty?.....

   Write in the space below any details you care to give regarding difficulties you may experience. (*Some persons shrink from casual conversation, either through shyness or through fear of wasting time. Others are too talkative or rambling. Some are curt. Many speak indistinctly, or in fragmentary remarks that are hard to interpret, etc.*)

   .....

   .....

2. **At social gatherings.**

   Do you use English in this way?.....Frequently?.....Do you find that this use of English presents any outstanding difficulty? .....

   Write in the space below any details you care to give regarding difficulties you may experience. (*In light general conversation some persons talk too much. Others hold back unduly, without contributing their part; many lack readiness in taking up or dropping a subject, in turning the conversation to give their companions an opening, etc.*)

   .....

   .....

Public Speaking

1. **Preparing a speech for a special occasion, a dinner, anniversary, etc.**

   Do you use English in this way?.....Frequently?.....Do you find that this use of English presents any outstanding difficulty? .....(Etc.)

Excerpt from a six-page questionnaire distributed by the National Council of Teachers of English, Committee on Place of English in American Life. A complete statement outlining the purpose of the inquiry and giving full directions for answering the questions accompanied the questionnaire.

(Example B)

SALARY SURVEY

Please give below in the columns provided, information regarding employees and salaries.

(a) Column 1, show total number of employees working at each occupation.
Note—*Exclude all officials, managers and assistant managers of departments.*

(b) Column 2, Occupation. Self-explanatory.

(c) Column 3, Maximum Rate. Give highest salary paid for this occupation.

(d) Column 4, Minimum Rate. Give starting rate for this class of work, or lowest salary paid for this occupation.

(e) Column 5, Average for All. This should be obtained by taking the total monthly payments to all employees of this occupation and dividing by the number of employees as shown in Column 1.

Note 1—*Give the salaries at monthly rates, excluding any bonuses, profit sharing, dividends, extra money, etc.*
Note 2—*If payroll is made up weekly, multiply by 4.3 and give monthly rates.*

No. Emp. (1)	Occupation (2)	Monthly Salary Max. (3)	Min. (4)	Avg. (5)	No. Emp. (1)	Occupation (2)	Monthly Salary Max. (3)	Min. (4)	Avg. (5)
	Addressograph Operator					Draftsmen			
	Advertising Men					Employment Department			
	Artists					Engineers—Construction			
	Auditors					Engineers—Electrical			
	Bank Tellers					Eng'rs—Elec. and Mechan.			
	Bookkeepers					Engineers—Mechanical			
	Buyers					Grounds Superintendent			
	Cafeteria Managers					Key Punch Operator			
	Cashiers					Laboratory Assistant			
	Chemists					Multigraph Operator			
	Clerks—Billing					Nurse—Graduate			
	Clerks—Chief					Nurse—Practical			
	Clerks—Cost					Office Boys			
	Clerks—File					Photographer			
	Clerks—Order					Stenog.—Head, or Priv. Sec.			
	Clerks—Rate					Stenographers			
	Clerks—Receiving					Technical Assistants			
	Clerks—Shipping					Telegraph Operator			
	Clerks—Stock					Telephone Operator			
	Clerks—Other					Timekeepers			
	Comptometer Operator					Tracers (Drafting)			
	Correspondent					Typists			
	Dictaphone Operator					Watchmen or Guards			
						Total Number Employees			

Questionnaire used by a large industrial concern to collect wage data from a number of companies in the community. The proposition of exchanging information of the kind was developed in personal interview previous to the use of this form and complimentary copies of the summarized results were distributed to all informants. (Size of questionnaire, 8½ x 11.)

APPENDIX F

PROBLEMS

Chapters II to VIII—Facts Required

1. The General Manager of a department store which has approximately 1,500 employes asks you to work up a program of personnel statistics for the company. Think the problem through and outline in some detail **(a)** the facts that you would want on basic records, and **(b)** the procedure you would follow in developing the service. Indicate **also** some data that would be required in special studies but would not be required on basic records.

2. In some instances, we should want the basic records to be so organized as to enable us to show the salesforce, for example, distributed by nationality. Indicate other associations that you would want to be able to get from the basic records.

3. In these department store facts why should we want absence data distributed by:

a. Married and single
b. Health class
c. Sex
d. Period of service
e. Occupational class
f. Nationality
g. Rates of pay
h. Length
i. Cause
j. Day of the week

Indicate in each of these cases wherein the analysis might suggest action which would be beneficial to the company. For example, if the distribution of absences by nationality should show that the absence rate was much higher for certain nationalities (other things being equal), it would be wise for the management to discontinue the employment of these types.

4. Why generally in a program of statistics should we want labor losses distributed by:

a. Day workers and night workers
b. Married and single
c. Health classes
d. Sex
e. Period of service
f. Occupational classes
g. Rates of pay
h. Nationality
i. Method of obtaining employe
j. Cause

5. The General Auditor of a telephone company who has some 1,000 employes reporting to him asks you to make a study of overtime in the organization for the last six months. What facts would you undertake to assemble? State specifically what you would expect to get out of each class of facts you set out to collect.

6. Suppose the analysis of #5 should indicate that as a usual practice the overtime in the department each week was equivalent to the full-time of three additional employes. What action would you consider for recommendation and further investigation, particularly if the company's practice allowed time-and-a-half or double-time for overtime?

7. Your company has established an Employes' Savings and Loan Plan. If you were managing this plan and were responsible for its success, what broad aspects of operation would you want to know currently. Assuming a definite type of organization, what detailed facts would you require to give you this information?

8. Outline in a general way the facts you would require in determining the cost to the company of **(a)** absences, **(b)** tardiness, **(c)** vacations, **(d)** holidays, and **(e)** labor losses.

9. What advantage is to be gained by the establishment of a proper standard for **(a)** labor losses, **(b)** absences, and **(c)** wages. Criticize the use of "last month" or "same month last year" as standards.

10. The company has carried on a safety campaign for the last year. What facts would you require to test the effectiveness of the results of the campaign? Why is not the number of accidents over the period or as compared with the year before sufficient. Assuming that the relative number of accidents is greater among new employes, what is the importance of analyzing the accidents by periods of service?

11. If you were asked to work up a statistical service covering the assets and liabilities and the revenues and expenses of your company what basic data would you want at hand? How does the handling of basic data in these classes differ from other facts, particularly personnel and operations?

12. You are asked to develop statistics showing a comparison of the common stock of your company with that of your competitors so far as the position in the market is concerned. What facts would you seek to obtain?

13. What data would you require to make a rough check on the return your company is realizing on its investment as compared with the return realized by your competitors or selected concerns?

14. What facts would you require to determine the sources of the capital represented by common and preferred stocks which the company has invested in the business?

15. You are asked to compare the yield on the company's bonds and preferred stocks with the market rate for money. What facts would you undertake to get?

16. Why should we want sales orders distributed by salesmen, days of the week, locality, size of order, and class of customer?

17. Your company wishes to compare the relative amount of business done with that done by selected competitors and the total done in the country. What data would be required to make the study?

18. In your particular business what facts would you require ready at hand in the statistical organization relative to wholesale prices, retail prices, production activity, sales activity generally throughout the country?

Chapter X—Internal Sources

1. Secure copies of the personnel records listed on pages 91 and 95, the the financial records listed on pages 96, 98 and 99 and operations records listed on pages 103, 105 and 106. The sources listed on page 129 Chapter X will be helpful in the selection of these records.

2. In the development of a program of statistics, from which of these records would you obtain (a) rates of pay, (b) absences, (c) number of employees, (d) sales orders, (e) shipments, (f) purchases, (g) units produced?

3. The "Application for Employment" is filled out by a prospective employe, and is kept on file in the personnel department, if one exists, or in operating offices otherwise. Following this example outline just how the following records might be used in the company:

a. Employe's Individual Record
b. Separation Form
c. Stock Register
d. Interest Detail Register
e. Sales Journal
f. Voucher Register
g. Bank Check Register
h. Collection Register
i. Stores Requisition
j. Contract Order
k. Creditor's Invoice
l. Production Schedule
m. Repair Order
n. Sales Order
o. Delivery Order

4. Where would you get the data on overtime called for in #5 under Chapters II-VIII above. Those called for in #7, 8, and 10 of the same group?

5. Where would you expect to secure data required in establishing a standard time period for the performance of a certain operation? Indicate other internal data that would ordinarily be obtained from an original source.

6. A department store wishes to determine the relative cost of delivery as among horse-drawn, gasoline, and electric motor vehicles. What data would you require to get the answer to this problem and where would you get them?

7. What facts would you require to determine the relative cost of (a) a cooperative system of delivery for a number of local retailers and (b) an independent system whereby each merchant maintained his own delivery force; and where would you get the facts?

8. From what specific company records would you obtain the following data:

a. Amount of raw material received by the company.
b. Amount of raw material consumed by the company.
c. Amount of supplies used by the company.

d. Factory deliveries, acceptances, and rejections.
e. Sales orders cancelled.
f. Sales orders taken distributed by localities.

Chapter XI—External Sources

1. Make a list of the most serviceable sources (original, primary or secondary) for the following groups of external facts:

a. Personnel— Wages
Absences
Labor turnover

b. Financial— Call money rates
Time money rates
Yield on stocks and bonds
Rate of return on investment earned by corporate industries generally
Bank clearings
Debits to individual accounts
Federal Reserve rediscount rate

c. Operations—Stock on hand of selected materials such as crude petroleum and gasoline, particularly those required by your own company
Total quantity produced in the country of the articles your company produces. The total amount imported. Exported. Total quantity produced in other countries
Agricultural facts
Mining facts
Manufacturing facts
Sales facts

2. What wage data have appeared in the *Monthly Labor Review* (U. S. Bureau of Labor Statistics) in the last six months? When was this publication begun? What other series of publications does the same Bureau publish?

3. What broad groups of data appear regularly in the *Monthly Labor Review?*

4. What classes of data appear regularly in the following publications?

a. The Commercial and Financial Chronicle
b. The Journal of Commerce
c. Dun's Review
d. Federal Reserve Bulletin
e. Information Bulletins—American Railway Association
f. Survey of Current Business—U. S. Department of Commerce

5. Tabulate all the data that are given about your local city in the U. S. Statistical Abstract. About capital invested in new enterprises. About the production of eggs. About war-risk insurance. About tobacco. Are all the data contained in this publication obtained from government sources? How frequently is the Abstract published?

6. Indicate sources of the following data:

a. Production of automobile passenger cars
b. Chain store sales
c. Department store sales
d. Sales of Studebaker, Chandler and Ford passenger cars
e. Production of locomotives
f. Unfilled orders of United States Steel Corporation
g. Data which would tend to reflect the market rate of pay for typists

Compare the figures available from different sources and explain any disparities between them.

7. How many wage earners were employed separately in the manufacture of matches, automobiles, boots and shoes, clocks, and linen goods, at April 15, 1914, in the United States?

8. What kinds of data are available from the following Commercial Service Organizations:

a. Bloomfield and Bloomfield, Boston, Mass.
b. National Bureau of Economic Research, New York City.
c. Bureau of Applied Economics, Washington, D. C.
d. Standard Statistics Company, New York City.
e. Babson Statistical Corporation, New York City.
f. Brookmire's Economic Service, New York City.

9. What kinds of data are available from your local Chamber of Commerce? Merchants' Association? Trade Associations? University Bureau of Research?

10. Select two or three sources which contain data of the same kind, and compare the separate groups of figures with respect to (a) their suitability, (b) their adequacy, and (c) their reliability for a selected, definitely specified purpose.

11. How would you go about determining whether data of a certain type were available from Federal Government sources? From the National Industrial Conference Board? The Federal Reserve Board? The National Automobile Chamber of Commerce?

Chapter XII—Collecting Data—General

1. In the "average weekly earnings" which appear currently in the *Industrial Bulletin*, published by the New York State Industrial Commissioner, what is the statistical unit? Are office employes included? Are both men and women employes included? Are part-time employes included? Are factories which are running only part-time included?

2. For purposes of making a statistical enumeration, how would you define a farm, a factory, an employe, a sale, a shipment, a delivery, a crime, a shoe store, an export, a bank, an industrial accident, a purchase, an absence, a price, or a business failure?

3. You are asked to make an analysis of sales for your company. What

would you take as the unit, orders taken, shipments made, bills issued, or goods paid for?

4. Select a half-dozen or more statistical studies from literature and establish definitely just what statistical units were used. Judging from the general subject of the study or from its title, what other units or modifications of the unit might have been used?

Chapter XIII—Method of Collecting Internal Facts

1. How would you undertake to obtain currently the basic record data which are called for in #1 in Chapters II-VIII above? If you should undertake to secure copies of company records which would furnish the required data, outline in detail the method you would employ to have these records reach you currently.

2. How would you arrange to secure the basic financial facts on the "Acquisition of Capital" listed in Chapter V? Why do you prefer the routine departmental report here and copies of company records for the data of #1 above?

3. How would you arrange to provide the use for your organization of the records listed in answer to problems #2 under Chapter X above?

4. Outline in detail the method of procedure you would follow in collecting the data called for in #5 of Chapter X above and prepare the forms or records you would use.

5. You are asked to collect data to serve as a basis for a study of the relative value to a retail coal company of horse-drawn and motor vehicles. Indicate the data you would require and how you would undertake to collect them.

6. What data would you require to determine the average length of a trip made by the machines of a taxi company and how would you collect the data?

Chapter XIV—Original Collection

1. Count the students who use the gymnasium facilities for the period of a week and classify them by the kind of activity engaged in.

2. How many motor vehicles pass the main entrance to the college campus in a day? How many pass the busiest corner in town?

3. Count the number of persons who enter the library in a day and classify them by hours of the day.

4. Solicit the approval and cooperation of the local traction company and count the number of passengers who board the cars at selected points of interest to the company.

5. With the cooperation of local retail merchants count the number of prospective customers who enter the stores and classify them as between men and women and those who buy and those who do not.

Chapters XV to XX—Collecting Data—Primary and Secondary

1. You are employed to make a detailed investigation as to sickness among workmen in a city of 100,000, particularly seeking to ascertain the proportion properly attributable to their employment. What would you consider the best method of getting this information?

a. Do you think insurance should be provided against occupational diseases? If so, should the government (State or Federal), the company, the labor organization, or the workman himself furnish the insurance?

2. You desire to determine the relative efficiency with which farming operations are conducted in Illinois and Kansas. What data would you require? How would you undertake to collect them?

a. It is contended that improvement of manufacturing methods has been far greater than the improvement of agricultural methods within recent years. If this is true, how do you account for it? Do you think it would be socially and economically beneficial to extend "big business" to the agricultural field?

3. As secretary of a labor union you are asked to conduct a questionnaire inquiry relative to unemployment in your organization throughout the country. Prepare a questionnaire, complete in every respect, to be used in this investigation.

a. What are the more reliable sources of unemployment statistics in the United States?
b. Discuss the reliability of Labor Unions as sources as compared with other private and with government sources.

4. A large industrial establishment whose employes are confined to one locality wishes to provide lunch-room service for its employes. In this connection, it is desired to collect, through the questionnaire method, pertinent information from other establishments who operate employes' lunch-room service. Prepare a schedule of questions to be sent to informants by mail, which call for data that have a significant bearing on the important problems involved in the study of the proposed lunch-room service.

a. From a broad economic point of view, do you think it advisable for an industrial establishment to provide lunch-room service for its employes? (Eliminate such special considerations as those where employes have no outside lunch-room facilities or are on duty during the night, etc.)

5. As advertising manager for a magazine, you wish to determine the potential effectiveness of your magazine as a medium for advertising a certain standard food product by assembling data which show the potential buyers among your subscribers? What data would you require? Prepare a questionnaire to be sent to your subscribers by mail to collect this information. What positive stimulation could you offer?

6. Select a problem of your own choosing and prepare a questionnaire to collect the data required in the solution of the problem. How many questionnaires will it be necessary for you to distribute?

7. Your company asks you to collect from a number of representative concerns in the community wage statistics for employes comparable with your own. Which method would you consider preferable, the questionnaire method or personal interview? Explain fully. Outline in detail a plan for collecting these data by personal interview.

8. From a selected group of retail merchants in your local community, collect data showing the number of sales made during each day for a period of one week, separately for men and women.

9. In cooperation with your local Chamber of Commerce, or Merchants Association, develop a plan of procedure and collect from local retailers data reflecting out-of-town sales distributed by localities.

10. What data appears in published sources relative to manufactured products in your local community? Check a part or all of these figures by collecting similar information directly from local manufacturers.

11. How does your own college or university rank among those of the state in number of students, number of members of the faculty, endowment, and years since it was founded?

12. Prepare a record to be used by the college library in collecting statistics relative to the use of the library facilities. Such a record should show for each book drawn from the library the title of the book, its general class; the name, sex, college class and other pertinent facts of the borrower; the hour of day, the length of time the book is held; and so forth. Provision should also be made for collecting facts relative to the use of the reference books and reading room.

13. Select from literature 8 or 10 questionnaires that have been used in the collection of statistical data and criticize them.

14. Prepare a questionnaire and distribute a small number (20 or 30) to selected informants, say, local merchants or those in neighboring cities. On the basis of the replies received what improvements can you make in the questionnaire? What do the replies suggest relative to the number of questionnaires that would be required for a more elaborate inquiry?

Chapter XXI—Editing

1. In cooperation with your local grocer prepare and distribute to his customers a questionnaire designed to collect information showing the preference for a specified list of breakfast foods. Carefully edit the replies, analyze and tabulate them, and with the assistance of the grocer interpret the results.

2. Secure copies of returned questionnaires from local sources—Chamber of Commerce, Merchants Association, or other agencies—or from other sources and examine them.

3. Carefully edit the questionnaires received in the inquiry suggested in #14 of Chapters XVI-XX above.

Chapter XXII—Types of Basic Records

1. Prepare forms to be used as ultimate primary basic records for the data you have outlined in connection with #1 under Chapters II-VIII above.

2. In maintaining a program of statistics in your company, what ultimate primary basic records would you require relative to:

a. Labor turnover, internal and external
b. Call-money rates
c. Time-money rates
d. Yield on stocks and bonds
e. Debits to individual accounts
f. Factory output
g. Transportation costs

3. What advantage is to be gained from the records developed under #2?

4. Develop typical secondary basic records for internal wage statistics, rates, factory output, receivables, and return on the company's investment. Why should you want these records in a practical statistical organization?

5. Secure copies of selected patented card records such as the Rand, Kardex, Acme, Index Visible, Library Bureau, Post-Index, and Globe-Wernicke products. Discuss the use of these cards as ultimate primary or secondary basic statistical records. Would they serve as ultimate primary records for wages of other concerns in the community? For interest rates? Stock and bond yields?

6. Secure copies of blank punch cards from the tabulating machine companies. Devise a card which when executed for each employe would serve as an ultimate primary basic record for the facts listed under "The Working Force" in Chapter III. Also for absences, labor turnover, rates of pay, sales, factory output, shipments, purchases, and other groups of facts.

Chapters XXIII to XXVI—Textual and Tabular Presentation

1. Present the data separately in tabular form which appear in the text on pages 241 and 242. Which requires the most space? Which is more readily inspected?

2. Find other examples of textual presentation in current periodicals, set the data up in tabular form, and compare the effectiveness of the two methods of presentation. Why do newspapers and periodicals, particularly the former, use the textual method so freely?

3. The following data are taken from a pamphlet prepared by the United Typothetæ of America, representing an analysis of 1922 printing costs developed from the "Annual Statements of Costs" obtained from members of the organization.

Rearrange the several classes of expense (a) in alphabetical order, (b) so that the general average in each group will be in ascending order, (c)

so that the group averages will be brought together. Discuss the relative effectiveness of the different arrangements.

PERCENTAGES OF 9-H CLASSIFIED EXPENSE ITEMS TO PLANT INVESTMENT
(From 1922 Composite Statement)

9-H Line No.		A	B	C	D	E	F	G	H	General Average
2	Rent	8.78	7.07	7.15	6.25	6.77	5.38	7.15	8.21	7.17
3	Insurance	1.11	1.77	1.56	1.68	1.54	1.58	1.24	.83	1.30
4	Taxes	.63	.77	1.21	1.24	1.32	1.14	1.06	.79	1.05
5	Depreciation	8.19	11.20	11.08	9.88	10.58	10.06	8.97	8.60	9.63
6	Interest on Plant Investment	5.14	5.93	6.39	5.94	6.25	5.91	6.75	5.25	5.90
	Fixed Expenses	23.85	26.74	27.39	24.99	26.46	24.07	25.17	23.68	25.05
8	Wages	47.04	59.02	65.21	68.02	80.10	74.24	85.66	88.25	79.24
9	Light	.60	.61	.51	.52	.54	.50	.47	.53	.53
10	Power	.89	1.04	1.33	1.24	1.27	1.05	1.73	1.97	1.54
11	Spoilage	.55	.57	.79	.78	.84	.83	1.00	1.37	1.01
12	Departments Direct Expense	2.32	3.06	4.00	5.10	6.24	8.57	7.95	14.19	8.77
	Current Expenses	51.40	64.30	71.84	75.66	88.99	85.19	96.81	106.31	91.09
20	Salaries	25.29	20.91	23.72	19.30	20.27	13.99	14.93	12.36	16.68
21	General Expense	3.23	3.81	4.41	4.19	4.98	3.72	6.62	3.59	4.43
22	Office Expense	2.10	2.44	2.25	2.12	1.89	2.09	1.96	1.17	1.76
23	Bad Debts	1.36	1.52	2.20	1.65	1.72	1.34	1.33	1.08	1.44
24	Shipping and Deliveries	1.28	1.47	1.47	1.31	1.53	1.07	1.56	.76	1.21
	General Expenses	33.26	30.15	34.05	28.57	30.39	22.21	26.40	18.96	25.52
27	Salaries and Commissions	5.26	10.11	11.86	9.85	12.26	11.12	12.84	5.24	9.42
28	General and Traveling Expense	.22	.99	1.61	1.40	2.35	2.06	6.21	2.35	2.56
29	Advertising	2.36	2.12	2.51	2.46	2.26	1.61	2.43	1.51	2.03
	Selling Expenses	7.84	13.22	15.98	13.71	16.87	14.79	21.48	9.10	14.01
36	Total 9-H Expenses	116.35	134.41	149.26	142.93	162.71	146.26	169.86	158.05	155.67

Classification represents plants with average Investment of, A—$5,287.50; B—$13,168.15; C—$23,589.48; D—$49,510.08; E—$83,630.02; F—$154,846.44; G—$211,126.32; H—$710,547.22, and General Average—$71,941.01.

(Source: United Typothetæ of America)

4. Select three or four tables from published sources and criticize them. Could you improve on them? How would you set the data up if you wanted the tables to serve different purposes?

5. Tabulate the data given on page 242 of Chapter XXIV showing the figures to the nearest thousand.

6. Taking the data from the U. S. Statistical Abstract, prepare a table showing the comparison between the domestic production and the importation (quantity and value) of the following commodities for the years 1916, 1917, 1918, 1919, and 1920: raw cotton, pig iron, iron ore, anthracite, bituminous coal, and cement. Set the table up in a number of different ways and select the one which you consider preferable, giving the reasons for your selection.

7. How would you tabulate for publication the data collected in the survey suggested in problem #3 of Chapters XVI-XX above?

Chapters XXVII to XXVIII—Graphics—General

1. Construct a single vertical scale, having one-half inch represent $100 of expenses. Locate the following expense items on this scale:

$300.00	$10.00
150.00	15.00
175.00	20.00
625.00	12.00
490.00	1.00
372.00	0.75
260.00	12.50

2. Using the scale of #1, construct a series of vertical bars to represent the expense items given in #1.

3. Construct a horizontal time scale projecting to the right from the foot of the vertical scale of #1, having one-half inch represent one month. Locate the following values in the field by means of these two scales, designating each one by a point:

	1923	1924
a	b	c
January	$125.00	$215.00
February	120.00	210.00
March	135.00	220.00
April	137.50	235.00
May	172.00	232.50
June	170.00	231.00
July	165.00	225.00
August	145.00	215.00
September	140.00	210.00
October	155.00	235.00
November	185.00	245.00
December	205.00	248.00

4. Connect the points established in #2 by a straight line, thus constructing a continuous broken line. Select a number of points on this broken line which lie between the plotted points and interpret their values on the vertical scale.

5. Prepare data sheets, complete in every detail, for the data given in problems #6 and 7 in Chapter XXIX below.

Chapter XXIX—Curve and Bar

1. On a single grid construct simple historical graphs from the two series given below, using the same scale for the two curves, and complete the chart for their proper presentation.

a. Is the "cost of living" a proper basis for determining rates of pay? If not, what is the proper basis? Justify your position.

Course of Average Weekly Earnings in New York State Factories

(Index numbers with June, 1914, as 100)

Month	1914	1915	1916	1917	1918	1919	1920	1921	1922	1923	1924
January	.....	98	107	120	132	181	209	217	192	206	219
February	.....	98	108	121	139	174	208	211	190	204	218
March	.....	100	110	124	147	175	219	212	193	212	222
April	.....	99	111	122	152	174	219	206	190	213	218
May	.....	100	112	127	157	175	224	204	194	218	217
June	100	101	113	128	161	177	227	202	196	219	214
July	99	100	111	127	164	182	224	199	195	217	213
August	99	102	114	129	167	188	226	200	198	214	216
September	98	101	117	134	176	196	226	197	202	216	221
October	97	105	118	136	176	192	228	193	202	218	217
November	97	106	119	139	170	200	226	191	205	218	218
December	99	106	122	139	183	207	223	196	208	220	222

(The Industrial Bulletin—The Industrial Commissioner of New York State)

Course of Retail Food Prices in the United States

(Index numbers with June, 1914, as 100. Derived from the relative prices published by the United States Bureau of Labor Statistics)

Month	1914	1915	1916	1917	1918	1919	1920	1921	1922	1923	1924
January	.....	104	108	129	162	187	203	174	143	145	151
February	.....	102	107	134	163	174	202	160	143	143	148
March	.....	99	108	134	156	177	202	158	140	143	145
April	.....	100	110	146	156	184	213	154	140	144	142
May	.....	101	110	153	160	187	218	146	140	144	142
June	100	101	113	154	164	186	221	145	142	145	144
July	103	101	112	147	169	192	221	149	143	148	145
August	108	101	114	151	173	194	209	157	140	147	146
September	108	102	119	155	180	190	205	155	141	151	148
October	106	104	122	159	183	190	200	155	144	152	150
November	106	105	127	157	185	194	195	154	146	153	152
December	106	106	127	159	189	199	180	152	148	152	153

(The Industrial Bulletin—The Industrial Commissioner of New York State)

b. Assuming an over-supply of labor and an increase in the cost of living, would rates of pay increase?

2. The Federal Department of Labor gives the following retail prices per ton of 2,000 lbs. of Pennsylvania coal for household use on the dates and at the places specified:

In January of each of the years 1915, 1916, 1917, 1918, 1919, 1920, 1921, the prices in Chicago were, in order: for the stove grade, $8.10, $8.10, $9.57, $10.35, $11.81, $12.59 and $15.91; and for the chestnut grade, $8.35, $8.35, $9.67, $10.39, $12.02, $12.69, and $16.03; while in New York City the prices for the stove grade were successively for the years indicated $7.14, $7.11, $8.50, $9.06, $10.76, $11.54 and $14.54; and for the chestnut grade $7.29, $7.25, $8.50, $9.08, $10.76, $11.60 and $14.54.

In July of the same years, prices of stove coal in Chicago were $7.90, $8.24, $9.58, $10.90, $12.20, $14.68, and $14.91 (May 15); and in New York City, $6.91, $7.39, $8.44, $9.30, $10.80, $13.07, and $13.13 (May 15); while the prices of chestnut coal in Chicago were $8.15, $8.49, $9.67, $10.98, $12.30, $14.79, and $15.06 (May 15); and in New York City $7.06, $7.42, $8.42, $9.29, $10.86, $13.07, and $13.12 (May 15).

Present the figures given above in tabular and graphic form by means of the bar chart.

a. Confining your consideration to prices alone, on the basis of the facts you have presented what economy in general should the consumer realize by purchasing the winter's supply of coal in July?

b. How do you account for the disparity in New York and Chicago prices? What further data would be necessary to show conclusively the importance of transportation costs as a factor in this disparity?

3. Point out the component parts in each of the charts constructed under #1 and #2 above. What are the different ways of associating the separate graphs with the data they each represent? Why did you make the selection you did?

4. Could the data of #1 have been presented by means of bars? Which is usually preferable in historical data having so many time periods? Under what conditions would you use bars in such a series?

5. Outline clearly the separate steps you followed in constructing the curve chart under #1 above. Why did you select the particular scales you used? How would it affect the appearance of the graphs if you should reduce the horizontal scale to one-fourth or one-half of its present value, leaving the vertical scale as it is? Or, vice versa?

6. Plot the following data in the form of straight line curves on a single grid.

FREIGHT CAR MOVEMENT

Average Loadings per Week. Selected Classes of Freight

1921, 1922, 1923

MONTHS	1921		1922		1923	
	Live Stock	Merchandise, L. C. L. and Miscellaneous	Live Stock	Merchandise, L. C. L. and Miscellaneous	Live Stock	Merchandise, L. C. L. and Miscellaneous
a	b	c	d	e	f	g
January	33,561	384,870	31,594	403,487	34,603	477,952
February	28,378	397,520	29,085	432,936	32,172	492,697
March	26,090	439,386	27,373	493,966	31,311	554,110
April	27,845	449,743	27,160	521,540	31,556	581,867
May	26,652	456,258	29,768	545,750	31,129	578,692
June	26,973	458,798	29,038	557,031	30,270	587,184
July	23,979	448,886	26,655	545,585	30,538	561,025
August	27,656	485,908	29,705	557,241	34,210	590,820
September	29,769	512,942	35,292	559,422	39,719	610,548
October	36,994	562,790	40,404	587,701	43,732	643,954
November	31,450	471,282	38,613	554,160	40,107	593,855
December	29,100	424,479	33,610	486,480	35,927	503,713
Average	29,037	457,739	31,525	520,441	34,607	564,701

(American Railway Assn.—Standard Statistics Company)

Plot the same curves using two vertical scales to bring them together. Plot the same curves, having a break in the scale to bring them together.

7. From the data on page 607 construct **(a)** a line-and-angle curve, **(b)** a stair-case curve, and **(c)** a rounded curve.

8. Smooth the line-and-angle curve of #7 by the following methods:

a. Free-hand

b. By taking the average for the entire series and for each half of the series.

c. By taking a three months' and a 12 months' moving average. Which gives the smoother curve? Which do you consider preferable?

d. By taking a 12 months' moving total.

Discuss the relative effectiveness of these different methods.

9. Prepare seven copies of the chart constructed under #1 above and associate the data with the graphs in the several positions listed on page 307. Discuss the relative effectiveness of the several positions.

10. Assemble a half-dozen or more graphs from current periodicals or other publications which in your opinion violate in an important way the principles of graphic construction. Point out some of the unfortunate results that might follow from their use.

(Data for Problem 7, page 606)

Value of Building and Engineering Contracts Awarded
(Data Represent 27 States)

(Unit $1,000,000)

Month	1910	1911	1912	1913	1914	1915	1916	1917	1918	1919	1920	1921	1922	1923
January	39.0	66.9	38.9	62.8	51.1	43.3	62.8	90.8	161.6	54.1	226.1	111.6	166.3	218.7
February	50.9	37.9	42.1	70.1	39.1	48.8	66.3	95.2	137.3	98.7	205.3	102.4	172.7	229.9
March	83.0	69.1	73.4	63.9	58.9	75.6	94.5	132.7	115.3	121.8	302.2	163.8	293.4	338.2
April	83.9	80.4	77.8	88.6	79.7	76.5	100.9	148.5	128.9	188.8	304.9	221.3	353.0	362.9
May	109.2	84.9	94.8	93.6	72.0	77.1	131.4	157.6	120.4	234.7	264.9	240.9	361.8	373.9
June	88.1	77.1	79.0	93.4	81.8	92.2	140.7	206.5	248.2	285.4	260.1	226.2	342.4	323.6
July	72.3	60.9	76.6	72.7	72.0	94.7	114.4	159.2	153.0	317.7	204.5	212.2	350.1	274.2
August	62.8	62.0	76.2	70.9	77.3	90.4	127.0	165.6	146.4	295.1	202.7	220.4	322.0	253.1
September	43.8	68.4	68.7	66.4	47.1	82.0	132.2	122.5	124.5	228.7	182.2	244.9	271.5	253.5
October	61.0	59.8	115.9	61.3	53.4	88.6	151.4	154.5	166.1	307.5	178.6	222.4	253.1	319.9
November	62.3	67.9	60.1	54.4	45.5	88.0	122.4	94.3	130.3	220.6	132.9	190.8	243.4	289.3
December	49.4	42.2	64.7	59.6	42.3	82.9	112.9	90.8	57.3	226.7	100.1	198.3	214.2	267.9
Total	805.7	777.5	868.2	857.7	720.2	940.1	1,356.9	1,618.2	1,689.3	2,579.8	2,564.5	2,355.2	3,343.8	3,505.1

(F. W. Dodge Corporation—Standard Statistics Company)

Chapter XXX—Forms of Curve and Bar Graph

1. Present the "General Averages" of the data given in problem #3 of Chapters XXIII-XXVI in the form of a series of columns.

2. Present the separate groups of expenses (Fixed Expenses, etc.) by means of columns, associating the bars in such way as to facilitate comparison **(a)** of the several classes of expenses for each group of plants, and **(b)** of each class of expenses for the several groups of plants.

3. Collect data representing the high and low prices of American Telephone and Telegraph Stock by month for the year 1924 and present the data in the form of a zone curve. What other kinds of data may be presented in this way?

4. From the following data construct **(a)** a band chart from the actual data, and **(b)** a 100% band chart. Consider the difference between the three columns and the total as "Miscellaneous."

STOCK OF FINISHED COTTON GOODS
By Months for 1923

Months	Stocks			
	Total *	White	Dyed	Printed
a	b	c	d	e
January	43,658	9,041	5,904	2,497
February	40,935	8,283	5,198	2,304
March	41,142	9,022	5,623	2,621
April	43,103	9,826	6,181	2,846
May	44,445	10,113	6,605	2,943
June	48,155	10,090	7,610	3,170
July	50,279	10,209	7,786	3,108
August	49,399	10,872	7,096	2,916
September	46,665	11,407	6,714	2,536
October	47,686	11,613	7,323	2,556
November	49,017	12,540	6,908	2,393
December	49,506	11,790	8,271	2,376

* Many plants were unable to give details as to classes of goods, so that the total column does not always agree with the sum of the component classes. In the computations called for the difference between the total and the sum of the other three columns should be considered as a miscellaneous column.
(Survey of Current Business)

5. Collect statistics of electric power production in the United States and prepare a table showing the simple and cumulative figures by months for the year 1923 and a 12 months' moving total also for each month. Plot these three series in form of Zee-Chart.

6. Secure copies of the Gantt chart from current periodicals and explain fully the data that it represents.

7. Devise a Gantt progress expense chart for a local merchant.

Chapter XXXI—Ratio Chart

(In these problems "semi-logarithmic" refers to a grid with the vertical scale logarithmic.)

1. Secure copies of commercial semi-logarithmic paper having the vertical scale logarithmic. Note that the principal graduations are designated by the single numbers 1, 2, 3, 4, etc. Locate the value 1.5 on the vertical scale. What is the relative distance of this point from 1 and 2? Why?

2. Locate the following values on the vertical scale:

2.5	9.5
3.6	7.75
1.8	5.25
6.9	4.1
8.2	8.15

3. Construct a logarithmetic scale consisting of a single cycle 8 inches long using the device noted on page 328 to transcribe the graduations from the slide rule or commercial paper.

4. Plot the data of problem #6 in Chapter XXIX above, on semi-logarithmetic paper. Compare these curves with the arithmetic curves originally constructed. Do the same for the data of problem #1 in Chapter XXIX above.

5. Plot the following data in straight line arithmetic and logarithmic curves:

ANNUAL AVERAGE EXPORT PRICES

Years	Anthracite, per long ton (In dollars)	Beef, per pound (In dollars)
a	b	c
1910	4.94	7.5
1911	4.99	8.7
1912	5.11	7.4
1913	5.30	9.6
1914	5.27	9.8
1915	5.27	10.6
1916	5.25	10.6
1917	5.56	11.6
1918	6.16	14.1
1919	7.22	20.2
1920	6.65	19.6

Discuss the difference in the appearance of the curves.

6. Plot each of the following series first on arithmetically ruled paper, second on logarithmically ruled paper:

(a)	(b)	(c)
10	10.0	100
12.5	12.5	200
15.6	15.0	400
19.5	17.5	800
24.4	20.0	1600
30.5	22.5	3200
38.1	25.0	6400
47.7	27.5	
59.6	30.0	

Chapter XXXII—Circle, Map and Pictures

1. From the following figures construct a series of circle graphs:

PERCENTAGE RATIOS OF FACTORY, COMMERCIAL, AND SELLING EXPENSES AND PROFIT TO SALES

	A	B	C	D	E	F	G	H	General Average
Materials Used	27.34	29.90	29.62	29.85	29.01	30.51	35.14	31.81	31.11
Stock Handling Expense	1.90	1.06	1.18	1.46	1.48	1.49	2.07	2.01	1.72
Factory Wages	27.61	27.42	26.91	28.12	30.65	33.17	30.44	31.93	30.47
Factory Expenses (CurrentandFixed)	13.98	13.12	12.27	12.47	12.40	14.72	11.55	14.19	13.08
Factory Cost	70.83	71.50	69.98	71.90	73.54	79.89	79.20	79.94	76.38
Selling Expense	4.75	6.53	7.13	6.05	6.88	7.13	8.21	3.54	5.78
Commercial Expense	21.56	15.95	16.26	13.94	13.99	12.12	11.38	8.92	11.98
Net (Trading) Profit	2.86	6.02	6.63	8.11	5.59	.86	1.21	7.60	5.86

(Source: United Typothetae of America)

Judging from the graphs which group of plant had the largest factory cost? Selling expense? Net profit? Has the construction of these graphs aided materially in bringing out the essential points of comparison in these figures?

2. Present the data of #1 above by means of a series of component bars. Are these more effective than the circles?

3. Present the data of #1 above by means of series of simple bars having the bars for a single vertical class grouped so as to facilitate comparison between the groups of plants. Thus, the 9 bars for "Materials Used" will be grouped together. How do these graphs compare with those of #1 and #2 in effectiveness in bringing out comparisons?

4. From page 3 of "Statistics of Income" (U. S. Treasury) for the year 1922, secure the personal income tax payments computed from the 1922 returns for the U. S. by states. Group these percentages into classes and construct a cross-hatched cartogram of the United States, representing these data.

5. Assemble from current periodicals a dozen or more picture and combination graphs and evaluate their effectiveness.

Chapter XXXIII—Analysis—General

1. In the table for problem #6 in Chapter XXIX above, total the data for the year 1921 vertically and horizontally and then summate the totals to obtain a grand total. Check the totals for the years 1922 and 1923 and for all three years in the same way.

2. Construct a table from the data of problem #6 of Chapter XXIX showing the cumulative figures for each year in addition to the consecutive items.

3. Compute a 3 months', 6 months' and 12 months' moving total from the data referred to in problem #1 above. Compute corresponding moving averages.

4. What is the ratio of the freshman class in college to the total student body? To the senior class? What is the ratio of fraternity to non-fraternity men?

5. What is the ratio of foreign born population in your city to the total population? Of women to men? Of qualified voters among the foreign born to those who are not qualified?

6. Secure the Balance Sheet and Income Statement of 10 or 12 industrial corporations and compute the following ratios from the detail contained therein:

a. Fixed capital assets to total assets.
b. Current assets to total assets.
c. Funded debt to property investment.
d. Current assets to current liabilities.
e. Common stock outstanding to property investment.
f. Common stock (par value) plus surplus to common stock.
g. Selling expense to total expense.
h. Net earnings to common stock.
i. Gross revenue to total investment.
j. Net earnings to total investment.

By reference to Bliss' "Financial and Operating Ratios in Management" discuss the significance of the results of your computations. Compute other ratios suggested in this book and explain what they show relative to the status of the business.

7. Taking the figures of #2 above to the nearest thousand, compute (a) consecutive percentages, (b) cumulative percentages, (c) distributive percentages. What is the advantage of each of these sets of percentages?

8. An employe's weekly pay was decreased 20% during the year 1921 and was increased 20% during the year 1923; there having been no other changes during the three years 1921, 1922 and 1923, how did his pay at the end of 1923 compare with that at the beginning of 1921?

9. If the wages of your factory workers were increased 10% during the year 1916, 20% in 1918, and 10% in 1920; and were decreased 40% in 1921; and then increased 10% again in 1923; how did the wages at the end of 1923 compare with the wages at the beginning of 1916?

10. The consecutive percentages for the sales of your company are as follows for the first 6 months of 1924:

January	—
February	85
March	130
April	105
May	101
June	95

How did the June sales compare with the January sales?

11. Your factory workers are working on a 48-hour week basis. You increase the hours to 50 but the weekly pay remains the same. What is the percentage change in the hourly rate?

12. The wages per hour are increased 10%, and the hours worked per week decreased 15%; what is the change in weekly wages?

13. The average price of cattle in January 1924 was $9.47 per 100 lbs. and in April of the same year it was $10.78 per 100 lbs. Enter the proper computations in the following sentences relative to these two prices:

a. The price increased —% during the period.
b. The price in April is —% of the price in January.
c. The price in January is —% of the price in April.
d. The price in January is —% less than the price in April.

What is the difference in the significance of these percentages?

14. From local concerns in your community collect data representing the average number of employes and the labor losses by months for a period of six months and compute the labor turnover by expressing the losses as a percentage of the average force. What other methods are there of computing labor turnover?

15. Find the percent increase in average rates of pay from January 1914 to 1921 of the following groups of employes, showing the increase separately for each of the three groups and for all employes:

	January 31, 1914		January 31, 1921	
	No. of Employes	Weekly Rate of Pay	No. of Employes	Weekly Rate of Pay
Clerks—Junior	1,100	$10.00	1,500	$15.00
Intermediate	800	12.00	1,000	18.00
Senior	400	14.00	500	21.00
Typists—Inexperienced	800	10.00	900	15.00
Experienced	550	12.00	600	18.00
Stenographers—Inexperienced	600	12.00	700	15.00
Experienced	250	16.00	300	25.00

16. The capital investment of your company is $50,000 and the plant investment is $100,000, what return must the company earn on its plant investment to provide a dividend payment of 10% on the capital investment?

Chapter XXXIV—Sampling

1. Collect a group of some 300 or 400 wage data and array them in ascending order. Take each fifth item in the series and compute the average rate. Begin with the second item and repeat and then with the third and the fourth. How do these averages differ from each other and from the average of the entire group? Repeat the process, taking each tenth item and each twentieth item. Compare these averages with those already determined.

2. You are asked to collect data on rates of pay from other companies comparable with those of your own company. How would you select the outside companies in order to get a representative picture of the wage rates in the community?

3. How does the U. S. Bureau of Labor Statistics select the items that enter into its index numbers of wholesale prices? Discuss fully.

4. Compare the number of items that enter into the wholesale price index series of the Bureau of Labor Statistics, Bradstreet, Dun, and the New York Times Annalist. Plot the several series on a single grid and compare the curves. How would you determine the items to use in the construction of such a series?

5. Locate statistical studies in current literature in which sampling has been used and critically examine the selection of the samples.

Chapter XXXV—Frequency

1. Collect data representing the rates of dividends paid on common stocks by 50 of the largest industrial corporations in the United States (based on capitalization at January, 1923) for the last 10 years.

   a. Arrange these data in order of size.
   b. Group them in classes with unity as class interval. In what group does a rate of 5.5 fall in this series? Plot these data in straight line frequency curve.
   c. Group the data in classes with 2 as a class interval. Plot this series in straight line frequency curve. How does it compare in regularity with the curve of the series in **(b)**?
   d. What has been the most frequent rate?

2. Basing your judgment on the analysis of #1 above, which would have netted the investor the highest return on his money, to have invested in the common stocks of these companies or to have placed it in the savings bank? (Disregard the factor of compounding the interest payments.)

3. Collect data and construct (a) a continuous frequency series, and (b) a discontinuous or discrete frequency series. By inspection of these two series explain fully the difference between the two and how the character of the series affects its analysis.

4. Construct a frequency table showing the accidents of your company distributed by period of service. Assuming that the accident rate is materially higher for new employes, explain how the accident rate for the company as a whole might decrease during a given period and yet accident conditions be really getting worse.

5. Secure from your local restaurateur the meal checks during each day for a week. Make a complete frequency analysis of these data and present them in tabular and graphic form. What practical benefit can the restaurateur get from your results in running his business?

6. Collect data showing accidents in the United States distributed by age groups, and plot the figures in a simple frequency curve. How does this curve differ from the form the frequency curve generally takes?

7. Collect data and construct a frequency table showing the entries consecutively and cumulatively in both ascending and descending order. Plot a "Less than" and a "More than" curve from these figures.

Chapters XXXVI to XL—Averages

1. Collect data representing the monthly production for 1924 of motor trucks, petroleum, tin, zinc, brick, cement, wheat, corn, apples, cotton, and shoes and compute the average monthly production.

2. Compare the average monthly imports and exports in quantity or value of coffee, cotton manufactures, sugar, tea, wheat, rice and wool.

3. There are 1,200 men employes and 800 women employes in your company. The average weekly rate of pay of the men is $40 and of the women $25. What is the average rate for the two groups combined?

4. In 1914 the average rate of pay of the women employes of the company was $12 per week, of the men $25 per week, and of the total $20 per week. In 1918, the average for the women was $16, for the men $30, and for the total $18. How do you account for the fact that the average for the total was less in 1918 than in 1914, although the average for the two groups separately increased materially during the same period?

5. The average weekly rate of pay in a company having 15,000 employes is $30. The average for the 1,500 office employes (all except the factory) is $40. What is the average weekly wage of the factory employes?

6. What would be the pay of each employe of the company if the total payroll were equally divided among the total number of employes?

7. Determine the *mean, mode* (group and value), *median* (term, group, and value), *lower quartile,* and the *upper quartile* of the following series:

Wages in Cents per Hour	Number of Workers
a	b
8 under 9	2
9 " 10	27
10 " 12	88
12 " 14	119
14 " 16	81
16 " 18	37
18 " 20	34
20 " 25	31
25 " 30	4
30 and over	0

a. In determining total payroll from the average per employe what type of average should be chosen? Discuss.

b. In considering compensation from the standpoint of the employes' welfare what type of average should be chosen? Discuss.

8. Referring to Table 2, on page 466, compute the mean, the median, and the modal rate of pay of the employes having 3 but less than 4 years of service. Do the same for other period of service groups.

9. Compute the *mean, median, and modal dividend rate* from the following data:

Dividend Rates on Common Stock

Dividend Rate	Frequency
a	b
3	3
4	4
5	15
6	25
7	4
8	6
9	2
10	7
Total	67

10. The following data represent the weights of packages of sugar sold at a chain grocery store during the week ending March 24, 1923:

TWO HUNDRED (200) RETAIL SALES OF SUGAR AT A CHAIN GROCERY STORE DURING THE WEEK ENDING MARCH 24, 1923

(Granulated Sugar—Loose)

Pounds	Sales
a	b
½	3
1	10
1½	1
2	5
3	7
3½	69
4	48
5	17
7	14
8	9
10	6
10½	1
12	4
15	3
20	1
25	2

Of what service should a study of this kind be to the grocer in determining the sizes of packages to prepare in advance?

11. Determine the modal price or prices from the restaurant checks analyzed in connection with problem #5 in Chapter XXXV. Does this suggest that it might be advisable for the manager to prepare combination meals at or about these prices?

12. The total sales of the company increased 75% during the 5 years 1916, 1917, 1918, 1919, and 1920. What was the average annual percentage increase?

13. What is the average size of a college class in your institution? What is the average number of books drawn from the library per day? What is the average number of class absences per week?

14. Secure the balance sheet statement for a selected company for the year 1923 or 1924. Compute the monthly balance in the property accounts by taking the sum of the balance at the beginning and end of the year and dividing by 2. Summate the balances at the end of each month and divide by 12. Compare the two results. Under what conditions may the average obtained by the first method be satisfactory for practical purposes?

15. Compute the quartile deviation, the average deviation, and the standard deviation from the arithmetic average of the wage series given under problem #7 above. Convert these to relative measurements or coefficients of dispersion.

16. Referring to Table 9, on page 393, compute the average rate of pay of the total number of employes. Change the class interval to 5 and then to 15 and compute the average in each case. How does the change in the class interval affect the average?

17. Compute the average and standard deviation from the data of #16 above and convert them to relative form. Also determine coefficients of skewness following the methods outlined in Chapter XL.

18. Compute the coefficient of skewness from the following data:

Income, Dollars	No.
a	b
0 – 200	10
200 – 300	70
300 – 400	560
400 – 500	530
500 – 600	280
600 – 700	150
700 – 800	120
800 – 900	37
900 – 1,000	22
1,000 – 1,100	12
1,100 – 1,200	8
1,200 – 1,300	5

Chapter XLI—Index Numbers

1. (1) On the left-hand half of a sheet of coordinate paper (approximately 8½ x 11) held horizontally, plot three curves representing the actual prices listed below, using the same scale for each curve.

Years	Actual Prices			Relative Prices		
	Cotton (lb.)	Coal (ton)	Steel (ton)	Cotton	Coal	Steel
a	b	c	d	e	f	g
1909	$0.12	$4.79	$24.62	156	133	114
1910	.15	4.81	25.38	195	134	117
1911	.13	4.81	21.46	168	134	100
1912	.12	5.03	22.38	148	140	104
1913	.13	5.06	25.79	165	148	120

(2) On the right-hand half of the same sheet, plot three curves representing the relative prices listed above, using the same scale for each curve although not necessarily the one used for actual prices.

a. Basing your conclusion upon the graphic representation of *actual prices*, which of the three commodities experienced the greatest fluctuation in price during the period under consideration?
b. Basing your conclusions upon the graphic representation of relative prices, which of the three commodities experienced the greatest fluctuation in price during the period under consideration?
c. In the comparison of the trend of prices what advantage is gained by the use of *relative* rather than *actual* price?

2. Compute an index series from the data of #1 above by the aggregative method.

3. Collect data representing bond sales on the New York Stock Exchange by months for the years 1922, 1923 and 1924. Using January 1922 as a base, convert the figures to relatives. What was the per cent increase or decrease approximately of the sales for June 1924 over January 1922?

4. If you have given the relatives of monthly sales for 1923 and the actual sales for September, how would you find the sales for other months?

5. From your local markets collect retail prices of sugar, meat, and eggs for the last 12 months. Convert each of these series of prices to relatives and plot the relatives together with the cost of living index of the Bureau of Labor Statistics on the same grid. How do the movements of local prices compare with that of the index? Is this a proper comparison? Discuss fully.

6. For your company, convert the expenses, revenues, and quantity of goods sold to relatives and plot on the same grid.

7. Develop a cost of living indices for the typical college student in your institution, following two separate methods as follows:

(a) By setting up a theoretical standard of requirements and then securing local costs for each.
(b) By collecting from a representative sample of students the amounts actually expended for different classes of items.

8. By inspection references indicated on pages 569 and 570, compare the general method of computation, the items included, and the averages and weights used in the computation of the following index series:

(1) Wholesale Prices—U. S. Bureau of Labor Statistics
R. G. Dun and Co.
Bradstreet's
(2) Cost of Living— U. S. Bureau of Labor Statistics
National Industrial Conference Board

9. Using the cost of living index of the Bureau of Labor Statistics convert the index of average weekly earnings of New York State factories to an index of real wages.

Chapter XLII—Correlation

1. Collect data showing liabilities of commercial failures to the nearest millions (R. G. Dun & Co.) and bank clearings in New York City to the

(Data for Problem 2, page 620)

Average Prices of Forty Bonds

Months	1915		1916		1917		1918		1919		1920		1921		1922		1923	
	High	Low	High	Low	High	Low	High	Low	High	Low	High	Low	High	Low	High	Low	High	Low
a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s
January	83.96	81.51	87.23	86.68	89.48	88.63	77.43	76.64	79.01	77.76	72.51	70.75	71.60	68.80	77.07	75.01	79.45	78.31
February	83.97	82.27	87.35	86.05	88.69	87.00	77.31	76.82	78.28	77.99	71.09	69.20	70.74	70.15	77.11	76.41	79.40	78.46
March	83.54	82.37	86.79	86.47	87.23	86.68	76.84	75.84	77.93	76.60	71.14	70.49	70.19	69.57	78.30	77.31	78.70	76.64
April	84.48	83.54	86.83	86.19	86.85	84.83	76.37	75.77	77.61	76.73	70.30	66.95	70.01	69.52	80.52	78.44	77.39	76.73
May	84.45	82.55	87.34	86.44	84.33	83.25	77.84	76.39	78.90	76.87	67.44	65.57	70.35	69.31	80.58	79.91	78.06	77.04
June	83.09	82.50	87.43	87.20	83.74	83.12	77.31	76.55	79.05	77.91	67.17	66.20	69.10	67.56	80.96	79.47	77.90	76.03
July	82.57	81.94	87.35	86.85	83.11	82.61	76.80	76.65	77.86	76.95	67.03	66.05	70.95	68.96	81.22	79.95	77.05	75.95
August	82.73	82.03	86.85	86.34	82.81	81.86	76.73	76.47	76.75	74.90	68.32	66.88	71.18	70.05	82.54	81.21	76.89	76.04
September	83.09	81.66	87.44	86.31	81.80	80.41	76.54	75.65	75.32	74.63	71.13	68.69	71.87	70.53	82.54	81.26	76.88	75.59
October	85.25	82.96	88.89	87.42	80.42	78.77	79.05	75.71	76.13	74.48	73.14	71.70	71.70	70.69	82.05	79.85	76.12	75.58
November	87.62	85.54	89.18	88.74	78.80	76.78	82.36	79.09	74.40	72.05	72.66	70.13	76.31	71.61	80.25	78.15	76.72	76.21
December	87.30	86.37	89.15	88.42	77.09	74.24	81.17	78.56	71.98	71.05	69.90	66.63	75.99	74.93	79.12	78.56	76.99	76.52
Year's ranges	87.62	81.51	89.18	86.05	89.48	74.24	82.36	75.65	79.05	71.05	73.14	65.57	76.31	67.56	82.54	75.01	79.45	75.58

(New York *Times*—Standard Statistics Company)

nearest billions, by months from the beginning of 1914 to the end of 1924. By means of the straight line curve plot these data on the same grid. Do the curves indicate correlation between these two series? Discuss fully.

a. What does a relatively large number of business failures indicate with respect to general business conditions? A relatively large volume of bank clearings?

2. Plot similar curves for the bond prices indicated on page 619 (a zone curve showing high and low) and interest rates listed below. Discuss the correlation between these two series suggested by the graphs.

AVERAGE OF HIGH AND LOW WEEKLY RATES OF 60-90-DAY CHOICE DOUBLE-NAME COMMERCIAL PAPER

Months	1913	1914	1915	1916	1917	1918	1919	1920	1921	1922	1923
a	b	c	d	e	f	g	h	i	j	k	l
January	4.93	4.53	3.81	3.13	3.55	5.58	5.19	6.00	7.83	4.88	4.63
February	4.91	3.84	3.75	3.13	4.09	5.69	5.19	6.41	7.75	4.88	4.69
March	5.75	3.88	3.38	3.13	4.13	5.88	5.38	6.68	7.63	4.78	5.00
April	5.53	3.73	3.66	3.13	4.28	5.90	5.38	6.81	7.55	4.60	5.13
May	5.36	3.88	3.72	3.13	4.83	5.88	5.38	7.16	6.88	4.25	5.13
June	5.88	3.84	3.65	3.63	5.00	5.88	5.53	7.72	6.63	4.05	4.88
July	6.06	4.40	3.25	3.97	4.68	5.88	5.42	7.84	6.28	3.94	4.94
August	6.00	6.34	3.53	3.73	4.81	5.94	5.38	8.00	6.00	3.91	5.03
September	5.78	6.70	3.25	3.38	5.19	6.00	5.38	8.00	5.90	4.25	5.16
October	5.69	6.44	3.22	3.38	5.38	6.00	5.38	8.00	5.65	4.38	5.13
November	5.56	5.50	2.98	3.50	5.47	5.97	5.50	7.94	5.13	4.63	5.09
December	5.68	4.35	3.13	3.91	5.50	5.86	5.88	7.88	5.13	4.63	4.98
Average	5.59	4.78	3.45	3.43	4.74	5.87	5.42	7.37	6.53	4.43	4.98

(*Commercial and Financial Chronicle*—Standard Statistics Company)

3. Collect data representing the prices of Bessemer Steel Billets (from *Iron Age*) and composite Pig Iron prices (from *Steel and Metal Digest*) for the years 1915 to 1924. Plot these two series in straight line curves on the same grid and note the tendency toward correlation between them. How does the correlation suggested here differ from that suggested in #1 and 2 above?

4. Collect data representing the production of automobiles, farm tractors, and farm lighting plants from the first years of their active distribution. Plot these in the form of straight-line curves and compare the general movements of the three. Plot the curves again, shifting the time scales so that the corresponding early years of each series will coincide. What does this latter arrangement suggest relative to the probable future course of farm tractors and farm lighting plants?

5. Collect data from the records in the department of physical education of your college showing heights and weights of students and prepare a correlation table from these data.

Chapter XLIII—The Time Series

1. By the method of moving averages compute the secular trend for the data on bank clearings in New York City for the years 1900 to 1924.

2. For the data on interest rates given in connection with problem #2 in Chapter XLII above, determine the secular trend by the method of moving averages. Use a 3-year period and a 5-year period and compare the results.

3. By the different methods noted in Chapter XLIII of the text, compute the seasonal variation from the data on "Value of Building and Engineering Contracts Awarded" given in connection with problem #7 in Chapter XXIX above.

4. Taking the figures to the nearest thousands, compute the secular trend of the figures given below on freight car loadings and list the deviations of the actual data from these ordinates. Compare the two graphically.

5. Collect data on department store sales (Federal Reserve Bulletin) for such years as may be given and correct the series for secular trend and seasonal variation and plot the resulting cyclical figures in the form of a rounded curve.

FREIGHT CAR MOVEMENT—AVERATE LOADINGS PER WEEK

Total All Classes of Freight

Months	1919	1920	1921	1922	1923
a	b	c	d	e	f
January	728,039	819,849	704,838	702,218	845,074
February	686,841	776,178	683,088	764,627	841,741
March	697,383	847,942	692,230	825,760	916,632
April	714,653	730,497	706,035	723,339	940,991
May	758,579	861,643	756,841	786,672	974,685
June	809,352	895,949	765,223	842,434	1,011,253
July	858,049	900,957	750,717	825,021	986,098
August	892,344	967,609	809,973	876,517	1,040,906
September	959,970	968,743	841,373	935,079	1,036,787
October	966,695	1,005,066	929,172	992,470	1,077,749
November	807,207	883,729	760,817	943,709	977,820
December	757,665	722,750	682,869	838,425	826,496
Average	803,065	865,076	756,931	838,023	956,353

(American Railway Assn.—Standard Statistics Company)

Chapter XLV—Mechanical Aids

TABULATING CARDS

1. Prepare a card to be used in the analysis of sales data to show sales by salesmen, districts, days of the week, classes of commodities, price, and classes of customers.

2. Prepare a card to show labor losses distributed by cause, sex, rate of pay, age, married or single, and period of service.

3. Secure copies of cards used by local companies and outline the data that can be obtained from them.

Slide Rule

4. Multiply 8 by 12. Find 8 on scale D and place the left index of C opposite it. Then, find 12 on scale C and read the number on scale D directly opposite it. This will be the product.

5. Multiply 4 by 8. Find 4 on scale D and proceed as in #1 except that the right index of scale C should be placed opposite the 4 on scale D. Why is not the left index of scale C used?

6. Multiply 57 by 23. Locate 57 on scale D, place the right index of scale C opposite it, and proceed as in #2 above. It will be noted that we cannot exactly read the last number as it falls just after the number 131. But since the product of the last two numbers is 21 we know that the last figure in the product we are seeking will be 1.

7. Make each of the above computations by finding the other number on scale D at the start.

8. Multiply 2 by 4 by 9. Find 2 on scale D, place the left index of scale C opposite it and place the runner at 4 on scale C. Then place the right index of scale C at the runner and opposite 9 on scale C read the product on scale D in the usual manner.

9. Compute the following multiplications:

21x13; 69x37; 71x11; 8x6x4; 9x8x42; 21x46x53; 125x8; 965x45.

10. Divide 9 by 3. Find 9 on scale D (note that the dividend is always located on scale D) and place 3 on scale C directly opposite it. Directly opposite the left index of scale C is 3, the quotient, on scale D.

11. Find the quotient of each set of figures given in #9 above.

12. The total weekly payroll of the company is $2,565 and the number of employes 150. What is the average weekly rate per employe?

13. A merchant purchases goods at 18½ cents a yard and wishes to determine at what price per yard he must sell it to make a profit of 15% on the cost price. Determine the selling price from the slide rule and explain the process.

GENERAL PROBLEMS

1. Outline the method of developing a statistical study to determine how much economy would result from decreasing accident compensation cases 5%.

2. Determine the equity of the common stockholders in 5 selected companies at December 31, 1924.

3. In cooperation with your local grocer, make a house-to-house canvass and develop a statistical study designed to indicate the demand for groceries in the territory which he serves.

4. a. How many golf balls were manufactured in the United States in 1923? How many exported? How many imported?

b. How much would it cost the company to decrease the working day one hour and not reduce wages? Discuss the method of procedure required in solving this problem.

c. If all the money the company is spending for employes' service activities were distributed among employes in the form of wages, how much would each get?

d. How much money has the company tied up in purchases, materials and supplies in stores, finished product on hand, and receivables overdue?

5. What should be the ratio of funded debt to capital stock in your company? Select a list of the soundest concerns in the same industrial class as your own and determine a standard by computing the ratios for each of these companies.

6. What is the cost in selling expense in your company to get $1 of revenue?

7. "In one year of inflation the nation's income passed $60,000,000,000. At that figure the average of the worker if the whole income went to workers, would be around $1,500 a year. That is the highest average wage that ever has been possible in this country. So the common talk about $1,800 or $2,000 or $2,200 a year as a general wage is moonshine. Neither the money nor the income is there to give it substance. We only raise false hopes when we talk such figures—sympathetic hopes but false ones." Selecting one year from the period 1919 to 1924, develop a complete statistical study to serve as a general check on the statement that $1,800 to $2,200 a year as a general wage is beyond realization.

a. From the broad economic point of view, what is the proper method of procedure in undertaking to increase the general average of real wages among workers in the United States?

8. "If all the net incomes above even $10,000 a year were seized, the present possessors turned out to grass, and the proceeds distributed among the forty odd millions of wage and salary workers, the gain of each worker would be about $2.20 per week." Confining your inquiry to one of the years 1920 to 1924, develop a complete statistical study to prove the truth or falsehood of this statement.

a. What is the relative importance of equal distribution of income and increased production, in seeking relief for the lower economic classes of the country?

b. How would you go about determining a real minimum wage at any given time?

9. Make a list of statements taken from your current reading and check their truth or falsehood or accuracy by the development of statistical studies.

10. How does the U. S. Bureau of Labor Statistics collect data to

develop its wholesale price index? Cost of living index? Outline and criticise the methods of determining the index from these data.

11. How many houses are electrically wired in your community? What is the demand for electric light bulbs?

12. In cooperation with your local Merchants' Association develop a statistical study indicating the total demand for men's ready-made clothing in your community.

13. What has been the real increase in Sears and Roebuck sales as compared with the money increase over the last five years?

14. "High prices and consequent high profits in any particular line of business attract capital and lead to greater activity in that line." Develop a statistical study to prove the truth or falsehood of this statement.

15. It is sometimes contended that the large industries of the country are making "enormous profits." Develop a sample study showing the rate of return realized by large industries of the country on the moneys invested. (Figures in "Statistics of Income" matched against figures in "Census of Manufactures" will serve as a broad check on the final results.)

16. It is sometimes said of a dividend-paying company that unless dividends are paid from surplus, the company is earning at least the dividend rate on its investment. What is wrong with this reasoning? Collect data relative to two or three companies in the United States where a relatively small return on the investment would be sufficient to provide for a relatively large dividend payment.

a. Discuss the relative soundness of over-capitalization and under-capitalization in an industrial concern.

17. "The bulk of the capital for the development of industries always has come from the industries themselves, and the larger the profits the faster the development has been. The year 1919 was one of high prices and high profits in the sugar industry and consequently the *production of sugar in the Caribbean section was greatly increased during that year.*" Confining your survey to the years 1914-1919 (both inclusive), develop a complete statistical study to prove the truth or falsehood of the statement in italics.

From the economic point of view,

a. What causes high prices of a commodity?
b. Which of the following policies would contribute most to the relief of an economic situation which was causing the high prices of a commodity:
   (1) *Government regulation* which holds the price of the commodity much below where it otherwise would be.
   (2) *Absence of regulation,* thus permitting the price to go as high as it will.

18. Develop a statistical study showing, in both actual rates and percentages, comparison in rates of pay or earnings in 1914 or 1915, 1920, and 1921, separately for office employes and factory workers.

(Depending upon the convenience of the data, the figures may be taken as averages for the years specified or as of a specific date—such as July 15, or January 1—for each of those years. Specific directions in regard to geographic territory, occupational classes, basis of compensation, etc., are purposely omitted.)

a. Which of the two general classes of employes considered realized the greatest increase in compensation from 1914 or 1915 to 1920? How do you account for this difference? Discuss fully.
b. What part, if any, do the questions of (1) overtime rates, and (2) additional compensation (bonus, profit-sharing, etc.) play in comparing rates of pay? Discuss each separately.

19. Outline the broad method of procedure in attacking the following problems from the statistical point of view:

a. Should men and women be paid equal compensation for equal work in your company?
b. To what extent has the capital stock of your company changed hands in the last 10 years?
c. Is this a good time to issue new securities?

20. A company established an Employes Benefit Plan in 1913, establishing a minimum old age pension in flat amount of $40 a month, and death benefits in flat amount of $1,000. In 1920 the company wished to change the flat amount relative to pensions and death benefits so that the allowances would be approximately equivalent in buying power to those of 1913. Determine the required amounts and explain your method fully.

21. What has been the "real" increase or decrease in the gross revenue and net profits of your company over the last 5 years as compared with the money increase? Outline the method of answering this question and discuss its importance in the management of the business?

22. Of the total amount of your kind of business being done in the country or other territorial section, what portion is your company doing? What portion should your company receive? Outline the broad methods of procedure in undertaking to solve these problems.

23. For industrial securities in general and for those of a selected group of concerns comparable to your company, how has the market value of securities compared with the par value for the last five or ten years?

24. What has been the average return to the individual investor in industrial enterprises in the United States so far as dividends on common and preferred stocks alone are concerned?

25. In general, what is the relative book value of capital stock obligations and the plant investment in industrial enterprises in the United States?

26. Of the fifty largest industrial concerns in the United States from the standpoint of the amount of capitalization, what per cent have paid no dividends to common stockholders in the last five years?

27. Which would have netted the greatest return in the last fifteen years, money invested in a reliable savings bank or money invested, equal

amounts, in the common stock of ten of the largest industrial concerns of the United States, the size of the concerns being determined by the capitalization? (Disregard the element of profit or loss through appreciation or depreciation in stocks and also the return due to compounding of interest or dividend payments.)

28. Outline the method of procedure in solving the following problems and work such of them up for local concerns as it is feasible to do:

a. What is the relative efficiency of workers of different nationalities in the several types of work in your company?
b. Assuming that the total payroll of your company remains constant, how much would the average compensation per employe be increased if all salaries over $5,000 a year were reduced to that figure, the excess being distributed equally among the employes receiving a compensation less than $5,000 a year?
c. What has been the trend of "real" wages in the company for the past 10 or 15 years?
d. What is the cost to the company of its labor turnover?

INDEX

A

Abscissae, 279
Absence Records, 91
  Form for, 94
Accounting (See also "Balance Sheet"; "Income Statement"; "Financial Reports")
  Statistics and, 515-520
Accounting Records (See "Financial Records")
Accuracy,
  Aimed at and attained, 138
  Importance of, 193-195
Acquisition of Capital (See "Financial Records—Acquisition of Capital")
Additional Compensation, Internal facts, 19
Additional Facts,
  External financial facts, 47
  External marketing facts, 70
  External operations facts, 78
  External personnel facts, 30
  Internal financial facts, 41
  Internal operations facts, 62
  Internal personnel facts, 24
  Treatment of, 14
Advertising,
  Historical internal facts, 58
  Incidental to collection of data, 159
  Internal facts, 56
Agricultural Facts, External production facts, 66
Analysis of Data, 353-492
  Aim of, 354
  By association, 87, 360
    In personnel field (Fig.), 361
  By classification,
    Extent to which classification should be carried, 359-360
    Frequency, 377-401
    Miscellaneous elements, 358-360
    Primary bases of, 355
    Qualitative, 355-356
    Quantitative, 356-357
    Time element, 357
  By computation, 360-362
    Checking total, 362
    Cumulative total, 362
    Moving total, 363
    Percentages, importance and use of, 364-368
    Rate defined, 364
Analysis of Data—*Continued*
    Ratios of various forms, 363-368
    Summations, 362-368
    Vital statistics terms, 364
    Weighted total, 362
  General, 353-368
  Meaning of, 353
  Need for, 353
  Processes of, 354
  Relation of to presentation of statistical data, 231
Antilogarithms, 571-576
Arithmetic Average (See "Averages—Arithmetic Average")
Arithmetic Curve, 318, 321
  Figs., 323, 324
Assembling Facts (See "Collection of Data")
Assets, Internal financial facts, 35
Association of Data (See "Analysis of Data—By Association")
Asymmetry (See "Averages—Skewness")
Attendance,
  Computed external facts, 28
  Internal facts, 17
Average Deviation, 440-442
Averages, 402-429 (See also "Index Numbers"; "Secular Trend"; "Time Series—Method of Moving Averages"; "Time Series—Method of Simple Averages")
  Arithmetic average or mean,
    Defined, 403, 404
    Effect of skewness on, 446
    Method of weighting, 407-408
    Nature and computation of, 406-407
    Short method of computing, 577-579
    "Step" method of computing (Table), 580
    Use of, 409-411
    Use of in index numbers, 459
    Weighting and the frequency series, 408-409
  "Assumed," 577-579
  Dispersion, 436-445
    Average deviation, 440-442
    Choice of measure, 443
    Coefficients of, 439, 444
    General conception, 436-437
    How to use measures of, 443
    Methods of measuring, 437
    Quartile deviation, 439-440
    Range, the, 437

AVERAGES—*Continued*
Relative, 444
Standard deviation, 442-443
Use of graphics, 443
Geometric, 404, 430-434
Computation of, 430-432
Definition and nature of, 429
Use of, 433-434
Use in index numbers, 459
Weighted, 432-433
Harmonic, 404, 434-435
Computation of, 434
Use of, 435
Meaning of, 402
Measuring variations from, 436
Median, 404, 412-419
Computation of, 412-413
Determined from frequency curve (Fig.), 417
Effect of skewness on, 446, 447
Graphical interpolation, 417-418
Grouped series, 415-417
Nature of, 412
Simple series, 413
Use of, 419
Mode, 404
Approximate, 422
"Approximate" distinguished from "true," 421
Continuous series, 424-426
Determination of, 422
Discontinuous series, 427
Effect of skewness on, 446
General conception, 420
Location in graphs, 427
Modal group, 423
Modal value, 424-426
Nature and characteristics of, 428
Selecting classes for data, 422
Series with more than one mode, 426-427
Tri-modal distribution (Fig.), 425
"True," 420-421
Use of, 428-429
Purpose of, 403
Reliability, 436-448
Skewness, 445-448
Methods of measuring, 446
Positive and negative, 445, 447
Relation between median and quartiles, 447
Symmetrical distribution of items rare, 445-446
Use of measures of, 447
Variation in relative values, 446
Types of, 404-405
AXES, 265, 278

B

BALANCE SHEET, Source of facts concerning use of and return on capital, 98
BAND-CURVE, 319-320, Figs., 320, 321
BANK CHECK REGISTER, 99
BANKING FACTS, External financial facts, 44
BAR GRAPHS (See "Graphics—Bar Graphs")
BASE LINE, 265
BASIC DATA,
Computed external facts, 27
Methods of maintaining, 217
Primary external, 27
BASIC RECORD FACTS,
Classification and association of facts, 230
External financial facts, 42-45
External operations facts, 63-75
External personnel facts, 25-29
Importance of proper, 217
Initial primary, 218-220
Accuracy of records, 213
Completeness of records, 211
Editing of, 211-216
Errors in, 215
Form of, 219
Uniformity of records, 212
Internal financial facts, 33-38
Internal operations facts, 48-57
Internal personnel facts, 16-22
Primary, 217-226
Routine reports as, 230
Secondary, 226-230
Detail to be included in, 227
Forms (with Figs.), 227, 228, 229
Special studies as, 230
Types of, 217-230
Ultimate primary, 220-226
Hollerith system, 223
Long-hand forms (with Figs.), 220, 221, 222, 223, 224
Patent card record systems, 223
BASIS OF STATISTICS (See Sources, Statistical)
BETTERMENT ORDER, 106
BOND REGISTER, 98
BONDS, Internal financial facts, 34
BONUSES, Internal facts, 19
BOOK, As example of difficulty of establishing statistical unit, 137
BUILDING FACTS, External operations facts, 72
BUREAU OF FEDERAL STATISTICS, Recommended, 118
BUSINESS, Use of statistics in, 7
BUSINESS CONCERNS, Statistical activities, 112, 121, 124
BUSINESS FACTS (See "Facts")
BUSINESS MANAGEMENT, Statistics and, 1

C

CAPITAL (See also "Financial Records—Acquisition of Capital")
Acquisition of, by other companies, 43
Acquisition of, internal facts, 33
Return on, in the community, 44

CAPITAL—*Continued*
Return on, internal facts, 37
Use of and return on, 98
Use of by other companies, 44
Use of, internal facts, 34
CAPITAL ASSETS, Internal financial facts, 35
CAPITAL LIABILITIES, Internal financial facts, 36
CAPTION, SCALES (See "Scale Caption")
CARDS (See also "Hollerith System")
For long-hand forms of ultimate primary basic records, 222
Fig., 224
Patent record systems, 223
Sorting machine for, 499-500
CASH JOURNAL, 99
CHART-FIELD, 265
CHARTS (See "Graphics")
CHECKING DATA, 150, 193, 588-592
CHECKING METHOD, For long-hand forms of ultimate, primary basic records, 221
Fig., 223
CHECKING TOTAL, THE, 362
CIRCLE GRAPHS (See "Graphics—Circle or Pie Graph")
CITY GOVERNMENTS (See "Municipal Governments")
CLASSIFICATION OF DATA (See "Analysis of Data—By Classification")
COEFFICIENT,
Of correlation, 468
Of correlation, Pearsonian, 580-585
COLLECTION OF DATA, 130-140 (See also "Business Facts—Distribution of"; "Personal Interviews"; "Questionnaire"; "Sources, Statistical")
Accuracy, 138
Advertising incidental to, 159
Amount of data required, 135
Checking, 150, 193
Defining the inquiry, 134
Estimating the cost, 136
From other business concerns, 197-202
Collected directly, 198-199
Commercial service organizations, 197
Form of data, 202
Permanent relations, 199-200
Personal interview, 199, 201
Presenting the inquiry, 199, 201
Regular collections, 200-202
Special collections, 202
Trade associations, 197
Importance of collection process, 130-132
Index numbers, 130
Internal facts, 140-148
Assembled by operating departments, 142-144
Assembled by statistical organization, 144-147
Duplicate financial records not needed, 147
COLLECTION OF DATA—*Continued*
Duplicate operating records preferable, 146
Duplicate personnel records obtainable, 147
Original collection, 147
Methods of, 140
Normal or representative conditions, 137
Original collection, problems of, 149-151
Method of making count, 150
Methods of recording observations, 150
Time, place and conditions of count, 149
Practical value of data, 135
Primary collection, 152-161
Aim of, 172
Confined to external field, 152
Establishing good faith, 155
Familiarizing informant with inquiry, 156
Immediate ends, 155
Methods used, 154
Selection of informants, 153
Stimulating favorable action from informants, 157-160
Ultimate ends, 154
Purpose, 133
Secondary collection, 203-210
General publications, 206-210
Government agencies, 203-205
Nature of, 203
Non-government agencies, 205-206
Selecting the type of collection, 138
Statistical unit, example of what constitutes, 136, 137
Survey of similar collections, 139
Thinking the problem through, 132-133
Types of collections, 132
COLLECTION REGISTER, 99
COLLECTION (See "Credits and Collections")
COMMERCIAL SERVICE ORGANIZATIONS,
Data furnished by, 123, 197, 206-207
Specific statistical service needed, 207-209
COMMISSIONS, Internal facts, 19
COMPANY INVOICE (Fig.), 109
COMPANY RECORDS (See also "Financial Records"; "Personnel Records")
Illustrative, 89
Location, form and designation of, 89
Primary, 86
Secondary, 86
Special, 87
Survey of existing, 90
COMPETITORS' SALES, External marketing facts, 69
COMPTOMETER, 494
COMPUTATION (See "Analysis of Data—By Computation")
COMPUTED FACTS, External personnel, 28

Computing Machines, 493-494
Conclusions (See "Interpretation of Statistics—Conclusions")
Construction Program, Internal, future estimates, 2
Consumption, External marketing facts, 69
Continuous Series, 380, 424-426
  Comparison of heights (Figs), 380, 381
Contract Order, 103
Coordinates, 279
Corporations, Statistical activities, 112, 121
Correlation, 464-470
  Coefficient of, 468
  Meaning of, 464-465
  Measuring, 465
    Computation of a coefficient, 468-469
    Correlation table, 466
    Use of graphs in (Figs.), 466-468
  Pearsonian coefficient of, 580-585
  Practical service from, 469
Cost,
  Of collecting statistical data, 136
  Of questionnaire, 168-169
  Statistical organization, 536-537
Cost Data,
  Marketing, internal facts, 55
  Of labor turnover, 19
  Production, internal facts, 52
  Purchasing, internal facts, 49
Cost and Expenses, Standards on, 40
Count (See "Collection of Data")
Creditor's Invoice, 103
Credits and Collections, Internal facts, 55
Cumulative Total, The, 362
Current Assets, Internal financial facts, 35
Current Liabilities, Internal financial facts, 36
Curves (See "Graphics—Curve Graphs"; "Frequency Classification—Frequency Curve")
Custom Trades, External marketing facts, 69
Cyclical Movements, 481-483

D

Data (See also "Analysis of Data"; "Basic Data"; "Business Facts"; "Checking Data"; "Collection of Data"; "Facts"; "Historical Data"; "Sources, Statistical")
  Entering of on table, 259
  Form of, 235
  Presentation of, 231-239 (See also "Textual Method")
    Meaning of, 232
    Method of treatment, 231
    Methods of, 233
    Purpose of, 233
    Relation of analysis to, 231
Data—*Continued*
    Selection of method of, 238
  Selection of, 234
  Should accompany graph, 306-308
  Woven into textual presentation, 241
Data Sheet, 282
  Fig., 283
Deductions, From wages, internal facts, 20
Defective Repairs, Rejected for, 106
Deferred Assets, Internal financial facts, 35
Delivery Records, 106
Demand, External marketing facts, 69
Derived Facts (See "Computed Facts")
Differentials, Internal facts, 20
Discontinuous, Series, 381, 427
  Number of sales by size (Fig.), 382
Dispersion (See "Averages—Dispersion")
Distributing Facts (See "Marketing Facts")
Divergence Curve (Fig.), 324
Dividend Book (with Fig.), 98
Draft, of statistical table, 257
Duo-Bar Charts, 318
Duplicate Records (See "Statistical Organization—Duplicate Records")

E

Earnings,
  Internal, future estimates of, 39
  Measures of, standards on, 40
Economic Statistics (See "Federal Government—Statistical Activities")
Editing, Basic records, 211-216
Educational Associations, Statistical activities, 126
Efficiency, U. S. Bureau of, 115, 120
Electric Key Punch, 498-499
Employes (See "Working Force")
Employes' Benefit, Authorization form, 95
Employes' Service Activities,
  Computed external facts, 28
  Internal facts, 20
Employment Conditions,
  Employes' service activities, internal facts, 20
  Hours of work, internal facts, 20
  Internal facts, 20
  Vacation practice, internal facts, 20
Employment Records, 91
  Changes, 91
    Form for, 92
Engagement Records, 91
Enumerators (See "Personal Interview"; "Questionnaire")
Equipment, Statistical organization, 536
Estimates (See "Future Estimates")
Expenditures,
  Income and, internal financial facts, 36
  Internal, future estimates of, 39

EXPLANATORY NOTES,
Footnotes distinguished from, 237
Statistical table, 262
EXPORTS, External marketing facts, 70
EXTERNAL FACTS (See "Sources, Statistical—for External Facts")
EXTERNAL FINANCIAL FACTS (See "Financial Facts, External")
EXTERNAL PERSONAL FACTS (See Personal Facts, External")

F

FACTS (See also "Additional Facts"; "Basic Record Facts"; "Data"; "Financial Facts"; "Historical Facts"; "Manufacturing Facts"; "Operation Facts"; "Personnel Facts"; "Sources, Statistical"; "Standard Facts")
Availability of, a requisite of scientific method, 4
Business, distribution of, 490-492
Detailed, the foundation of a statistical service, 11
Fundamental classification of, 12
Required in program of statistics, 11-15
FEDERAL GOVERNMENT (See also "Government Agencies")
Statistical activities, 113-121
Department lists of publications, 118-119
Practical method for studying, 114
Report on, 115
Serviceableness of, 119
Units contributing (Fig.), 116-117
FEDERAL STATISTICS BUREAU, Recommended, 118
FINANCIAL FACTS,
External, 42-47
Additional facts, 47
Basic record facts, 42-45
Data, 42-45
Financial facts of other companies, 43
Future estimates, 46
General facts, 44
Historical facts, 45
Money market, 42
Return on capital in the community, 44
Security market, 42
Standard facts, 46
Internal, 33-41
Acquisition of capital, 33
Additional facts, 41
Assets, 35
Basic record facts, 33-38
Bonds, 34
Expenditures, 37
Future estimates, 38
Historical facts, 38
Income and expenditure, 36
Liabilities, 36
Mortgages, 34
FINANCIAL FACTS—*Continued*
Notes, 34
Returns on capital, 37
Standard facts, 40
Stocks, 34
Use of capital, 34
FINANCIAL PLAN, Future estimates of, 38
FINANCIAL RECORDS, 95-99
Acquisition of capital, 96-99
Bond register, 98
Dividend book, 98
Interest detail register, 98
Mortgage register, 98
Stock register, 98
Stock subscription blank, 96
Stock transfer book, 96
Dividend book (Fig.), 98
Duplicates not required, 147
Importance of secondary, 95
Of other companies, 112
Standardized, 95
Use of and return on capital, statistics for, 98-99
Sales journal (Fig.), 100
Voucher register (Fig.), 101
FINANCIAL RELATIONSHIP, Standards on, 41
FINANCIAL STATISTICS,
Defined, 33
Figs., 557-562
FOOTNOTES,
Explanatory notes distinguished from, 237
Purpose of, 238
Statistical table, 262
FREQUENCY CLASSIFICATION, 377-401
Bell-shaped distribution, 393-394
Characteristics and use of frequency analysis, 398-401
Class limits, 389
Indication of (Figs.), 390, 391
Classes imposed by nature of data and purpose, 388
Continuous and discontinuous series, 380-382
Comparison of heights (Figs.), 380 381
Number of sales by size (Fig.), 382
Cumulative series, 396-398
(Figs.), 397, 399
Entering the data, 392
Form of frequency series, 392
Frequency curve, 378-379
Median determined from (Fig.), 417
Method of plotting (Fig.), 379
Meaning of, 377-378
Miscellaneous forms, 394-395
J-shaped distribution (Figs.), 394 395
Selecting the classes, 386-388
Minimum number, 387
Required degree of accuracy, 387
Simple series, 395-396
Expressed in percentile form (Fig.), 396

FREQUENCY CLASSIFICATION—*Continued*
Statistical array, 384-386
Five cent groups (Fig.), 385
Simple arrangement, one cent interval (Fig.), 384
Two cent groups (Fig.), 385
Unsystematic arrangement (Fig.), 383
Steps in analyzing frequency series, 383-384
FUTURE ESTIMATES,
External financial facts, 46
External operations facts, 76
External personnel facts, 30
Internal financial facts, 38
Internal operations facts, 59
Internal personnel facts, 22
Treatment of, 13

G

GANTT CHARTS, 316-317
Fig., 317
GENERAL CASH BOOK, 99
GENERAL CONSUMPTION FACTS,
External marketing facts, 71
GENERAL FACTS, External financial facts, 44
GENERAL INFORMATION (See "Additional Facts")
GEOMETRIC AVERAGE (See "Averages—Geometric Averages")
GLOSSARY, 539-547
GOVERNMENT. AGENCIES (See also "Federal Government"; "Municipal Governments")
Collection of data from, 203-205
GRAPH, THE, 266
Titles of, 235
GRAPHICS, 264
Abscissae, 279
Aim of, 266
Axes, 265, 278
Bar graphs, 280-322
Base line, 265
Broad bar (Fig.), 309
Combination of, series of simple bars (with Figs.), 313-315
Component bar (Fig.), 281
Duo-bar charts, 318
Forms of, 311
Gantt charts (with Fig.), 316-317
Historical data and, 315
Progress charts (with Fig.), 316
Series of component bars (with Figs.), 311-313
Series of simple bars, 311
Simple bar (Fig.), 280
Thin line bars (Fig.), 308
Variations in space between bars (Figs.), 309, 312, 314
Chart-field, 265
Circle or pie graph, 339-341
Figs., 340, 341
GRAPHICS—*Continued*
Classes of, 270
Column graph (See "Graphics—Bar Graph")
Combination graphs, 349
Figs., 348, 349, 350, 351
Component parts of chart, 265
Confused chart (Fig.), 267
Construction of, 271-279
Summary, 278-279
Coordinate lines, 277, 279
Correlation, determination of, 466
Curve-graphs, 280-325
Arithmetic curve (with Figs.), 318, 321
Band-curve (with Figs.), 319-321
Line-and-angle, 300
Logarithmic curve, 321
Miscellaneous forms, 321
Rectilinear or stair-cased, 301
Scale or type of curve (Fig.), 286
Smoothed, 301, 302-305
Zone-curve (with Fig.), 318-319
Data sheet, 282
Fig., 283
Dispersion portrayed by, 443
Explanatory notes, 305
Fundamental fact of, 271
Graduations on the simple scale (with Figs.), 276-277
Graph, the, 266
Designations, 306
Parts of (Fig.), 267
Titles of, 235
Grid, 265
Key or legend, 266
Logarithmic or ratio graphs, 326-337
Comparative value, 336-337
Construction of ratio scale (with Figs.), 326-327
Distinctive features (with Figs.), 328-332
Forms, 327
How to construct, 332
How to read, 333
How the ratio graph is used (with Figs.), 335
Maps, 342-345
Color and cross-hatch maps (with Figs.), 343-345
Pin and ribbon maps, 345
Meaning of, 264
Number of curves cannot be arbitrarily limited (Figs.), 268, 269
Ordinates, 279
Origin, 265, 278
Picture-graphs, 345-348
Figs., 346, 347
Plotted points, 266
Positive and negative values, 275
Indicated by direction of scales (Figs.), 275, 276
Quadrants, 279
Scale, 265-266, 279
Breaking of (Fig.), 295

GRAPHICS—*Continued*
Caption, 266
Designations, 279
Seasonal variation shown in, 480-481
Single scale (with Figs.), 272-275
Vertical scale (with Fig.), 274
Steps in construction, 284
Bringing curves into juxtaposition (with Figs.), 295-296, 297, 299
Connecting the points, 300-303
Data should accompany graph, 306-308
Equal distance represents equal values, principle violated (Fig.), 287
Final inspection, 306
Graph designations, 305
Locating the points, 300
Method of bringing out monthly figures (Fig.), 287
Methods of smoothing curves (with Figs.), 305
Ratio between scales, 292-295
Scale as a point (Fig.), 286
Scale as a space or point, 289
Scale as a space (Fig.), 290
Scale caption, 296-299
Scales, the, 285-289
Selecting the grid or field, 285
Series and scales, 291-292
Title, explanatory notes, footnotes, 305
Zero line shown (Fig.), 290
Zero point, 289
Summary of principle forms, 322-325
Thinking the problem through, 282-284
Time element placed on horizontal scale, 291
Title, 265, 305
Two scales in association, 275
Vertical and horizontal scales associated (Fig.), 278
Zero point, 265
GRID, 265

H

HARMONIC AVERAGE (See "Averages—Harmonic Average")
HISTORICAL DATA,
External financial facts, 45
External operations facts, 75
External personnel facts, 29
Graphic presentation of, 315
Internal financial facts, 38
Internal operations facts, 57
Internal personnel facts, 22
Treatment of, 13
HOLLERITH SYSTEM, 494-498, 501,502
Cards (Figs.), 494, 495, 497
Cooperation in use of, 521
For ultimate primary basic records, 223
HOURS OF WORK, Internal facts, 20

I

IDLE EQUIPMENT REPORT (Fig.), 105
IDLE LABOR REPORT, 105
IDLE MACHINERY, Internal facts, 51
IMPORTS, External marketing facts, 71
INCOME AND EXPENDITURE, Internal financial facts, 36
INCOME STATEMENT, Source of facts concerning use and return on capital, 98
INDEX, SEASONAL, 587
INDEX NUMBERS, 449-463
Aggregates of actual data, 453
Average of relatives, 453, 458, 459
Chain, 459
Collecting the data, 130, 457-458
Extensive use of, 449
"Ideal formula," 451
Interpretation of, 461
Numbers indicate trend only (Fig.), 462
Methods of computation, 451-453
Nature and purpose of, 449-450
Pertinent facts relative to a selected group of (Fig.), 569
Relatives of aggregates, 457, 458
Selecting the average, 459
Selecting the base, 458
Selecting the items, 457
Selecting the weights, 460-461
INDUSTRIAL MANAGEMENT,
Scientific tendency of, 2
INDUSTRIAL ORGANIZATION,
Statistical activities, 112, 121, 124
INFORMANTS (See also "Questionnaire")
Difficulties encountered with, 155
Selection of, 153
Should be familiarized with inquiry, 156
Stimulating favorable action from, 157
By distributing samples, 159
Through desire for personal benefits, 158
Through scientific interest, 159
Through a spirit of cooperation, 160
INITIAL PRIMARY BASIC RECORDS (See "Basic Records, Initial Primary")
INQUIRERS (See "Personal Interview")
INSPECTION, FINAL REPORT, 106
INTEREST DETAIL REGISTER, 98
INTERNAL FACTS (See "Collection of Data—Internal Facts"; "Personnel Facts—Internal"; "Sources, Statistical—For Internal Facts")
INTERPRETATION OF STATISTICS, 484-492
Aim of, 486
Conclusions, 488
Meaning of, 489
Distributing business facts, 490-491
Importance of, 489-490
Meaning of, 484-485
Review of principles and problems of management, 487
INTERVIEW (See "Personal Interview")
INVESTIGATOR (See "Personal Interview")

K

KEY, THE, 266
KEY PUNCH, ELECTRIC, 498-499

L

LABOR CONDITIONS, External personnel facts, 31
LABOR REPORT, IDLE, 105
LABOR TURNOVER,
Computed external facts, 28
Internal facts, 18
LEAST SQUARES, METHOD OF, 585
LEGEND, THE, 266
LIABILITIES, Internal financial facts, 36
LINK RELATIVES, METHOD OF, 587
LOGARITHMIC CURVE, 321
LOGARITHMIC GRAPHS (See "Graphics—Logarithmic Graphs")
LOGARITHMS,
Antilogarithms, 571-576
Composition of, 503
How to determine, 504
Properties of, 504
Theory of, 503
To find, 571-572
LONG-HAND FORMS,
For ultimate primary basic records, 221
Figs., 221, 222, 223, 224

M

MACHINERY, IDLE, Internal facts, 51
MACHINERY, COMPUTING, 493-494
MAINTENANCE AND REPAIRS,
Internal facts, 51
MANUFACTURING FACTS,
External production facts, 67
MANUFACTURING ORDER (Fig.), 105
MAPS (See "Graphics—Maps")
MARKETING, Standards on, internal, 61
MARKETING FACTS, Internal operations facts, 53-57
MARKETING RECORDS, 106-109
MARKETING REQUIREMENTS,
Historical internal facts, 58
Internal future estimates, 59
MATERIAL,
Available, external purchasing facts, 64
Consumed, external purchasing facts, 64
Distributed, external purchasing facts, 64
Produced, external purchasing facts, 65
MEAN, THE (See "Averages—Arithmetic Average or Mean")
MECHANICAL AIDS IN STATISTICAL WORK, 493-514
Centralization of statistical organization saves duplication of, 526
Computing machines, 493-494
Hollerith system, 494-498, 501-502
Cards (Figs.), 494, 495, 497
Need for, 493
MECHANICAL AIDS—*Continued*
Punching machine, 496-498
Electric key punch (Fig.), 498
Slide rule, 502-513
Construction and operation (Figs.), 505-507
How to read, 507-510
Locating decimal points, 512
Logarithms, 503-504
Practical operation (Fig.), 511-512
Use of, 512-513
Sorting machine, 499-500
Fig., 500
Summary and conclusion, 513-514
Tabulating machine, 500-501
Fig., 500
Verifying machine, 499
MEDIAN (See "Averages—Median")
METHOD OF LEAST SQUARES, 585
METHOD OF LINK RELATIVES, 587
MINING FACTS, External production facts, 67
MODE (See "Averages—Mode")
MONEY FACTS, External financial facts, 45
MONEY MARKET, Facts on, 42
MORTGAGE REGISTER, 98
MORTGAGES, Internal financial facts, 34
"MOUNTAIN" CURVE (Fig.), 324
MOVING TOTAL, THE, 363
MUNICIPAL GOVERNMENT, Statistical, activities, 122

N

NORMAL CONDITIONS, Must be reflected in data selected, 137
NORMAL FACTS (See "Standard Facts")
NOTES, EXPLANATORY,
Statistical table, 262
NOTES, Internal financial facts, 34
NOTES AND BILLS PAYABLE JOURNAL, 99
NOTES AND BILLS RECEIVABLE JOURNAL, 99

O

OPERATING RECORDS,
Assembled by operating department, 142-144
By statistical organization, 144-145
Extra copies for statistical organization, 145-147, 148
Marketing, 106-109
Company invoice (Fig.), 109
Sales order (Fig.), 107
Shipping order (Fig.), 108
Producing, 105-106
Delivery records, 106
Process records, 105
Production order records, 105
Purchasing, 99-104
Ordering records, 103
Receiving records, 103

Operating Records—*Continued*
Requisitioning records, 103
Stores records, 103
Statistics (Figs), 563-566
Operations Facts,
External, 63-79
Additional facts, 78
Basic records facts, 63-75
Future estimates, 76
Historical facts, 75
Miscellaneous facts, 72-75
Production facts, 65-68
Purchasing facts, general, 65
Purchasing facts, special classes, 64
Standard facts, 77
Internal, 48-62
Additional facts, 62
Basic record facts, 48-57
Future estimates, 59
Historical facts, 57
Marketing facts, 53-57
Producing facts, 50
Purchasing facts, 48
Standard facts, 60
Ordering Records, 103
Purchase order (Fig.), 102
Ordinates, 279
Origin, 265, 278
Original Enumeration, 128
Other Companies, Financial facts of, 43
Overtime Payments, Internal facts, 19

P

Patent Card Record Systems,
Ultimate primary basic records, 223
Pay Roll,
Form, 91, 93
Journal, 99
Total, internal facts, 20
Pearsonian Coefficient of Correlation, 580-585
Percentages (See also "Link Relatives, Method of")
Application or use of, 366
Consecutive (Fig.), 366, 367
Cumulative (Fig.), 366, 367
Distributive (Fig.), 366, 367
Miscellaneous, 367
Relatives, 367
Distinguished from "rates," 364
Importance of, 365
Periodicals, General,
Statistical material in, 127, 206
Furnished by commercial service organizations, 206-207
Methods of examination, 209-210
Personal Interview, 182-196
Accuracy, uniformity and completeness, 193-195
Conducting the interview, 186-188
Establishing contact, 185
Expensive in time, 195
In collection of data from other business concerns, 199
Personal Interview—*Continued*
Instructions to inquirers, 183-184
Number of inquirers, 182
Qualifications of inquirer, 182
Recording the data, 188
By the informant, 190-191
By the inquirer, 188-190
By inquirer and informant in collaboration, 191-193
Supplementing questionnaire, 169-170
Personnel, Of statistical organization, 531-536
Personnel Facts,
External, 25-32
Additional facts, 30
Attendance facts, 28
Basic data, 26-29
Basic record facts, 25-29
Comparison of with internal, 25
Computed facts, 28
Employes' service activities, 28
Future estimates, 30
Groups considered not individuals, 25
Historical facts, 29
Labor turnover facts, 28
Scope of, 26
Standard facts, 30
Wage facts, 28
Internal, 16-24
Additional facts, 24
Basic record facts, 16-21
Comparison of with external, 25
Employment conditions, 20
Future estimates, 22
Historical facts, 21
Standard on normal facts, 23
Working force, 16-19
Wages, 19
Personnel Records, 91-95
Analysis of data, 360
Fig., 361
Duplicates needed, 147
Engagement, 91
Of other companies, 112
Separation, form for, 95
Service, 91
Employes' benefit authorization form, 95
Employes' record changes form, 92
Payroll form, 91, 93
Report of absence (form), 94
Tardiness report, 95
Statistics (Figs.), 550-556
Pictures (See "Graphics—Picture Graphs")
Pie Graphs (See "Graphics—Circle or Pie Graphs")
Plant Equipment, Internal facts, 52
Plotted Points, 266
Population Facts,
External marketing facts, 71
External operations facts, 74
Price Data, External operations facts, 72
Primary Collection (See "Collection of Data")

Probability, Theory of, 372-373
Problems, 582-593
Process and Final Inspection Report, 106
Process Records, 105
  Idle equipment report (Fig.), 105
Producing, Standards on, internal, 61
Producing Records, 105-106
Producing Requirements,
  Historical internal facts, 58
  Internal future estimates, 59
Production Facts,
  External, 65-68
  Internal operations facts, 50
Production Order Records,
  Manufacturing order (Fig.), 105
  Production order (Fig.), 105
Production Orders,
  Internal facts, 50
  Requisition, 105
Production Schedule, 105
Profit-Sharing Payments, Internal facts, 20
Progress Charts, 316
  Fig., 316
Progress Curve (Fig.), 323
Public Finance Facts, External financial facts, 45
Publications, General,
  Statistical material in, 127, 206
    Furnished by commercial service organizations, 206-207
    Methods of examination, 209-210
Punching Machines, 496
  Electric key punch (Fig.), 498
    With duplicating attachment (Fig.), 498
Purchase Journal, 99
Purchase Order, 103
  Requisition, 103
  Schedule, 103
Purchased Materials Received Report, 103
Purchasing, Standards on, internal, 61
Purchasing Facts,
  External,
    General facts, 65
    Special classes, 64
  Internal operations facts, 48-50
Purchasing Records, 99-104
Purchasing Requirements,
  Historical internal facts, 58
  Internal future estimates, 59

Q

Quadrants, 279
Qualitative Classification, 355-356
Quantitative Classification, 356-357
Quartile Deviation, 439-440
Quartiles, 447
Questionnaire (See also "Personal Interview")
  Comparative cost, 168-169
Questionnaire—*Continued*
  Construction of, 163, 172-181
    Check questions, 176
    Components, 174
    Definitions, 178
    Directions, 178-179
    Explanatory notes, 177
    Mechanics, 179
    Position of explanations, definitions and directions, 176
    Questions, 174-176
    Thinking the problem through, 173
    Work sheets, 172-173
  Distribution of, 163, 164
  Examples, 591, 592
  Follow-up, 164
  Meaning of, 162
  Number of returns, 164-167
  Obtaining best results from, 170-171
  Quality of returns, 167-168
  Supplemented by personal interview, 169-170
Quotations, Request for, 103

R

Range, The, 437
Rates,
  Defined, 364
  Distinguished from percentages, 364
Rates of Pay, Internal facts, 19
Ratio, 363-368
Ratio Graphs (See "Graphics—Logarithmic Graphs")
Receiving Records, 103
  Materials received report (Fig.), 103
Recording Observations, 150
Records (See "Basic Records"; "Company Records"; "Financial Records"; "Operating Records"; "Receiving Records"; "Requisitioning Records")
Rejected for Defective Repairs, 106
Rejected for Scrap, 106
Reliability of Statistical Data, 84-85
Repair Order, 106
Repairs (See "Maintenance and Repairs")
Reporters (See "Informants")
Reports (See also "Business Facts—Distribution of"; "Routine Reports")
  Suggestions on (with Figs.), 549-570
Request for Quotations, 103
Requisitioning Records, 103
Retail Trade, External marketing facts, 70
Return on Capital (See "Capital, Return on")
Returns (See "Questionnaire")
Revenues, Internal, future estimates of, 39
Routine Reports,
  As basic records, 230
Rulings of Statistical Table, 260

S

Sales Journal (Fig.), 100
Sales Orders (Fig.), 107
  Internal facts, 54
Sales Return and Allowance, 99
Salvaged Material, Internal facts, 51
Sampling, Statistical, 369-376
  Conscious selection, 373
  Ideal sample, 371
  In practical work, 371
  Limitations, 376
  Mathematical tests, 375
  Meaning of, 369-371
  Practical procedure, 373
  Random selection, 374
  Size of sample, 374
  Theory of probability, 372-373
  Treatment limited in scope, 369
Scale Caption, 266, 296-299
Scales (See "Graphics—Scales")
Seasonal Index, 587
Seasonal Variation, 478-481
Secular Trend, 478-481, 585
Security Market, Facts on, 42
Selection of Informants, 153
Separation Form, 95
Service Records, 91
Shipping Order (Fig.), 108
Shop Deliveries, Internal facts, 50
Shop Deliveries of By-Products, Internal facts, 51
Skewness (See "Averages—Skewness")
Slide Rule, 502-513 (See also "Logarithms")
  Construction and operation, 505-507
    Figs., 506-507
  How to read, 507-510
  Locating decimal points, 512
  Practical operation, 511-512
    Fig., 511
  Use of, 512-513
"Slip-Stick" (See "Slide Rule")
Sorting Machine (Fig.), 499-500
Sources Statistical (See also "Basic Data"; "Collecting of Data")
  Adequacy of data, 84
  Classification of, 80-82
  Evaluation of, 82-83
  For external facts, 110-129
    Commercial service organizations, 123
    Educational associations, 126
    Evaluation of, 111
    Federal government, 113, 115-121
    General publications, 127
    Industrial organizations, 124
    Miscellaneous agencies, 127
    Municipal governments, 122
    Non-governmental agencies, 122
    Original enumeration, 128
    Other business concerns, 112
    State governments, 121
    University bureaus, 126
  For internal facts, 86-109
    Association of data, 87
Sources Statistical—*Continued*
    Evaluation of sources, 88
    Financial records, 95-99
    Illustrative records, 89
    Location, form and designation of company records, 89
    Marketing records, 106
    Operating records, 99-106
    Personnel records, 91-95
    Primary company records, 86
    Producing records, 105
    Purchasing records, 99
    Secondary company records, 86
    Special company records, 87
    Survey of existing company records, 90
  General, 80-85
  Reliability of data, 84-85
  Suitability of data, 83
Space Classification (See "Analysis of Data—By classification—Space element")
Special Studies, As basic records, 230
Standard Deviation, 442-443
Standard Facts,
  External financial facts, 46
  External operations facts, 77
  External personnel facts, 30
  Internal financial facts, 40
  Internal operations facts, 60
  Internal personnel facts, 23
  Treatment of, 13
State Government, Statistical activities, 121
Statements,
  Operating department, 143
  Suggestions on (with Figs.), 549-570
Statistical Array (See "Frequency Classification—Statistical Array")
Statistical Data (See "Data")
Statistical Organization, 521-538
  Centralized vs. decentralized, 522-528
    Centralized, 523-527
    Decentralized, 527-528
  Cost, 536-537
  Duplicate records,
    Financial, 147
    Operating, 145, 147, 148
    Personnel, 147
  Equipment, 536
  Gradual development preferable, 538
  Nature and function of, 4
  Personnel, 531-536
    Executive statistician, 531-533
    Number of employes, 536
    Staff statistician, 533-534
    Statistical clerks, 535
    Statistical experts, 534
  Place in the company, 521
  Practical form, 528-530
  Relation to accounting department, 515-520
    Cooperation, 519-520
    "Fact" department, 517-518
    Methods distinguished, 516-517

Statistical Organization—*Continued*
Practical considerations, 517
Purpose of accounting, 515-516
Type of data, 516
Use of operating copies, 144
Statistical Research Organization, Need for, 207-209
Statistical Schedule (See "Questionnaire")
Statistical Tables, Construction of, 251-263
Alphabetical arrangement, 255
Casual relationship, 255
Chronological arrangement, 255
Defining the purpose, 252
Distinctive lines, 261
Distinctive type, 258
Entering data, 259
Explanatory notes, 262
Final inspection for accuracy and completeness, 262
Footnotes, 262
Methods of emphasis, 258
Numbering the columns, 260
Order of arrangement, 254-256, 259
Order of size, 255
Position, 258
Putting table in final shape, 260
Relative position, basis of arrangement (with Table), 255, 256
Rough working draft, 257
Rulings, 260
Selecting the main and sub-classes, 253
Size of items the basis of arrangement (Table), 254
Special markings, 259
Steps in process, 252
Titles of, 235, 262
Statistical Unit, 136
Statisticians, 531-535 (See also "Statistical Organization")
Statistics (See also "Analysis of Data"; "Averages"; "Correlation"; "Index Numbers"; "Informants"; "Mechanical Aids")
Business facts (See under that heading)
Business management and, 1
Checking data (See under that heading)
Collection of data (See under that heading)
Defined, 6
Federal government (See under that heading)
Financial (See under that heading)
Glossary of terms, 539-547
Graphics (See under that heading)
Interpretation of (See under that heading)
Need for, 11
Operating records (See under that heading)
Personnel records (See under that heading)
Problems, 582-593
Statistics—*Continued*
Program of, facts required in, 11-15
Program of, for small and medium-sized companies, 15
Program of, typical, 14
Rapid development of, 1
Sampling (See under that heading)
Sources (See under that heading)
Suggested reports and statements (with Figs.), 549-570
Time series (See under that heading)
Use of in business, 7
Stock Register, 98
Stock Subscription Blank, 96
Stock Transfer Book, 96
Form for, 97
Stocks, Internal financial facts, 34
Stores Records, 103
Stores Requisition, 103
Stores Return, 103
Summations, 362-368
Surplus, Internal financial facts, 36
Symmetry (See "Averages-Skewness")

T

Tables (See "Statistical Tables")
Tabular Method, 246-251
Aim of tabulation, 248
Meaning of, 246
Tables analytical, 250
Tables component parts of, 247
Fig., 248
Tables, informative, 249, 250
Fig., 250
Tables, types of, 249
Tabulating Machine (Fig.), 500-501
Tally Method,
For long-hand forms of ultimate primary basic records, 221
Fig., 222
Tardiness Report, 95
Textual Method, 240-245
Advantages of, 243
Criticism of, 243
Data woven into, 241
Defined, 240
Forms of presentation, 240
Solid running text, 242
Theory of Probability, 372-373
Time Element (See also "Analysis of Data—By Classification—Time Element")
Placed on horizontal scale, 291
Time Series, 471-483
Cyclical movements, 481-483
Eliminating secular trend, 481
Eliminating secular trend and seasonal variation, 482
Nature of the problem, 481
Fundamental elements, 471-472
Seasonal variation, 478-481
Graphic methods, 480-481
Method of link relatives, 480
Method of moving averages, 479

Time Series—*Continued*
Method of simple averages, 478
Secular trend,
Free hand method, 472
Method of least squares, 478
Method of simple averages, 472
Moving averages (Fig.), 473
Title,
Graphics, 235, 265
Of statistical tables and graphs, 235, 262
Tables, 235, 262
Trade Associations, Collection of data from, 197
Transportation and Communication Facts, External operations facts, 74
Turnover (See also "Labor Turnover")
Standards on, 40
Type, Statistical tables, 258

U

Unfilled Orders, Internal facts, 51
Unit, Statistical, 136
United States Bureau of Efficiency, 115, 120
University Bureaus, Statistical activities, 126

V

Vacation Practice, Internal facts, 20
Variation, Seasonal, 588
Verifying Machine, 499
Vertical Scale, 274, 277
Vital Statistics, Terms used in, 364
Voucher Register (Fig.), 101

W

Wages,
Additional compensation, internal facts on, 19
Computed external facts, 28
Deductions, internal facts, 20
Internal facts, 19
Overtime payments, internal facts, 19
Rates of pay, internal facts, 19
Total payroll, internal facts, 20
Weighted Total, 362
Wholesale Trade, External marketing facts, 70
Working Force,
Attendance internal facts, 17
Internal facts, 16-19
Labor turnover internal facts, 18
Number and character of, internal facts, 16

Z

Zee-Curves (Fig.), 323
Zero Point, 265, 289
Fig., 290
Zone-Curve, 318-319
Fig., 319

MANAGEMENT

AND ADMINISTRATION

IN MANUFACTURING INDUSTRIES

A useful journal for manufacturers. Covers all essential problems of profitable manufacturing under highly competitive conditions. Articles are written for practical managers by men with practical experience. Material is not popularized, not inspirational, but consists of actual cases with actual results, organized where possible into tables, charts, and diagrams. Meets standards of progressive establishments. Users need not be technically trained. Monthly, 50c per copy, $5.00 per year, domestic or foreign. Sample copy on request.

BOOK PUBLICATIONS

REFERENCE AND EDUCATIONAL

(A Year Book containing a subject index to the information presented in these volumes will be sent on request.)

Affelder. CONTRACTORS' ACCOUNTING PRACTICE. Embraces the whole field of construction accounting for small and large contracting firms, with owners' and architect's records. 386 pp. $5.00.

Alford (Editor). MANAGEMENT'S HANDBOOK. Practical information completely covering management in manufacturing companies of production, financial, marketing, and interrelated activities. 32 sections contributed by 35 authorities. 1607 pp. $7.50.

American Bankers Association. SCHOOL SAVINGS BANKING. The method approved by the Savings Bank Division of the American Bankers Association, for installing and operating a system. 174 pp. $1.25.

American Engineering Council. INDUSTRIAL COAL—PURCHASE, DELIVERY, AND STORAGE. Report of nation-wide investigation by eminent engineers into practical methods of purchasing, delivery, and storage of coal for industrial use. 419 pp. $5.00.

Associated Knit Underwear Manufacturers of America. COST CONTROL FOR KNIT UNDERWEAR FACTORIES. Specific formulation of procedure based on a 5 years' investigation throughout the industry. 259 pp. $5.00.

Babb. BUSINESS LAW. Elementary textbook, sound in law, well arranged, and written in simple language. Hundreds of cases definitely cited. 489 pp. $1.75.

Bacas. AUDITING PRACTICE SET. Practice material to illustrate procedure in making an actual audit, preparation of working papers and finished report. To be used with "*Auditing Principles,*" by Montgomery and Staub. $2.00.

Baker. THE PREPARATION OF REPORTS—ENGINEERING, SCIENTIFIC, ADMINISTRATIVE. Report writing for engineers, scientists, industrial and commercial executives, and research workers. 468 pp. $3.50.

Belding. BUSINESS CORRESPONDENCE AND PROCEDURE. A brief textbook in letter-writing which makes real to high school students the viewpoint of the business world. 383 pp. $1.40.

Bell. ACCOUNTANTS' REPORTS. The best professional practice in form and method of presentation. Over 70 typical statements. For private as well as public accountants. 386 pp. $6.00.

Bliss. FINANCIAL AND OPERATING RATIOS IN MANAGEMENT. Shows how to use 18 such ratios and turnovers, and presents for comparison statistical compilations of standard ratios for 43 industries. 396 pp. $6.00.

Bliss. MANAGEMENT THROUGH ACCOUNTS. A thorough practical discussion of co-ordinating the accounting methods and policies of a business organization with the requirements of executive management. For both accountants and managers. 851 pp. $6.00.

Bloomfield. LABOR MAINTENANCE. A handbook of industrial service work embodying the procedure of well-known concerns. 530 pp. $5.00.

Boettiger. EMPLOYEE WELFARE WORK. A systematic study. Examples of success here and abroad and historical account from medieval guilds to the present. 301 pp. $2.50.

Brewster. LEGAL ASPECTS OF CREDIT. Law of sales and creditors' rights and remedies, and procedure to recover funds on involved accounts. Includes principles of credit granting. 549 pp. $5.00.

Brisco. RETAIL SALESMANSHIP. A high-school textbook. Vol. I—Basic principles in simple terms. Vol. II—A Source Book of readings, provides a background. Vol. I, "Retail Salesmanship," 278 pp. $1.50. Vol. II, "Source Book." 352 pp. $2.00.

Brown. EXPORT ADVERTISING. Organization and operation of advertising departments of concerns in foreign trade. Specific details with charts and forms. 342 pp. $4.00.

Burdick. ADVERTISING TO RETAILERS. Covers advertising by manufacturer to retail trade to win co-operation in marketing. 308 pp. $3.50.

Burnham. BUILDING YOUR OWN BUSINESS. How to start; discovery and test of idea; planning, organizing, financing, and operating. Unvarnished records of over 50 successes, with details. 282 pp. $2.75.

Cartmell. STORES AND MATERIALS CONTROL. Subject treated in full from viewpoint of production executive, including procurement by manufacturing and by purchase. 459 pp. $4.50.

Chapman. FISCAL FUNCTIONS OF THE FEDERAL RESERVE BANKS. U. S. Treasury operations through the Reserve Banks, including funding of government's floating debt and transfer of funds. 213 pp. $2.50.

Clapp-Lee. LANGUAGE FOR MEN OF AFFAIRS. Vol. I—Practical training for improving command of language in conversation and public speaking. Vol. II—Analysis of working principles governing all writing required of men in business and professional life. Vol. I, "Talking Business," 526 pp. $4.00. Vol. II, "Business Writing," 611 pp. $4.00.

Clark. THE GANTT CHART. Important new type of graphic chart for management and control. 157 pp. $2.50.

Clemen. THE AMERICAN LIVESTOCK AND MEAT INDUSTRY. Complete account. Colonial conditions: growth of Chicago plants; refrigeration; chemistry and by-products; operating methods; stock raising and marketing; labor relations, federal regulation. 872 pp. $6.00.

Cole. COST ACCOUNTING FOR INSTITUTIONS. A comprehensive system for hospitals, hotels, schools, and other public and private institutions. 248 pp. $3.50.

Conyngton (H. R.). FINANCING AN ENTERPRISE. A practical guide in securing capital for enterprises large or small. 3 vols., 667 pp. $7.00.

Conyngton (T.). BUSINESS LAW. For everyday use of business men. Includes 1,000 questions from C. P. A. and other examinations and 79 legal forms. 842 pp. $6.00.

Conyngton (T.)-Bergh. BUSINESS LAW. A textbook for schools of business, presenting more briefly for students main topics of Conyngton's "Business Law." 431 pp. $3.00.

Conyngton (T.)-Bennett-Pinkerton. CORPORATION PROCEDURE. Standard manual of the legal, financial, and accounting procedure involved in organization and management of corporations. 1689 pp. $10.00.

Conyngton (T.)-Knapp-Pinkerton. WILLS, ESTATES, AND TRUSTS. A manual of law, accounting, and procedure for executors, administrators, and trustees in making wills and settling estates. 825 pp. $6.00.

Cook. FINANCING EXPORTS AND IMPORTS. Covers foreign trade financing methods from standpoint of banker and foreign trade merchant. 218 pp. $2.50.

Cooke. PUBLIC UTILITY REGULATION. Present-day problems of state control of power and light, transit, gas, water, and telephones, discussed by 17 leading experts. 310 pp. $3.25.

Cowdrick. INDUSTRIAL HISTORY OF THE UNITED STATES. A general history in terms of economic, business, and labor factors. Particular attention to recent conditions. 414 pp. $2.75.

Daggett. THE HISTORY OF THE SOUTHERN PACIFIC. Building the system; Huntington, Stanford, and associates; financing; reorganizations; relations with government; transcontinental rate structure; Harriman control; divorce from Union Pacific. 470 pp. $5.00.

de Haas. FOREIGN TRADE ORGANIZATION. A comprehensive description of institutions and methods of international trade the world over. Useful bibliography. 378 pp. $3.00.

Dewey-Shugrue. BANKING AND CREDIT. Thorough discussion from viewpoint of business men dealing with banks. 506 pp. $3.00.

Dewing. CORPORATION FINANCE. College textbook based on the author's "Financial Policy of Corporations." (See next page.) 457 pp. $2.75.

Dewing. THE FINANCIAL POLICY OF CORPORATIONS. Exhaustive treatment of financial structure and problems. Citations and commentaries illustrating every point from experience of railroads, utilities, and industrials. 5 vols., 953 pp. $12.00.

Dickinson. ACCOUNTING PRACTICE AND PROCEDURE. Special points in advanced accounting; corporation accounting and finance; cost work; responsibility of accountant in public practice. 309 pp. $4.00.

Dohr. COST ACCOUNTING THEORY AND PRACTICE. This work co-ordinates cost work with general accounting. Stress on purposes and form of cost records, and their use by management. 631 pp. $4.00.

Dolch. MANUAL OF BUSINESS LETTER WRITING. A concise office desk book for quick reference. Definite suggestions. 100 sample letters illustrate layout, wording, and spirit. Practice Material. 357 pp. $2.50.

Dooley. APPLIED SCIENCE FOR METAL AND WOOD WORKERS. For vocational training classes, and apprentices and shop workers. 559 pp. $2.00.

Douglas. MERCHANDISING STUDIES OF THE STATES. Illustrates a method of merchandising investigation through analysis of topography, climate, and spirit of the people. 178 pp. $2.00.

Droppers. OUTLINES OF ECONOMIC HISTORY IN THE NINETEENTH CENTURY. Concise review of European and American development in a definite historical period. 286 pp. $2.50.

Edgerton–Bartholomew. BUSINESS MATHEMATICS. A textbook covering the simpler calculations required in ordinary business practice. 305 pp. $2.00.

Eggleston. MUNICIPAL ACCOUNTING. For large and small cities. Based on New York City methods. 456 pp. $5.00.

Esquerré. APPLIED THEORY OF ACCOUNTS. Application of sound theory to everyday accounting problems, including account classification, financial statements, partnership and corporation accounting. 575 pp. $5.00.

Esquerré. PRACTICAL ACCOUNTING PROBLEMS. For preparation for examinations or practice work. Part I contains 20, and Part II, 30 problems. Facsimile typewritten form showing presentation. Part I, 353 pp., $10.00. Part II, 356 pp., $10.00.

Farnham. AMERICA vs. EUROPE IN INDUSTRY. Detailed comparison of foreign and domestic business methods; administrative and executive control, factory buildings, machines, labor; shop government; profit-sharing. 492 pp. $4.00.

Ferris–Collins. SALESMANSHIP. A thorough and logically organized discussion, covering present-day conditions. 414 pp. $3.00.

Fordham–Tingley. ORGANIZATION AND BUDGETARY CONTROL IN MANUFACTURING. Essential principles and detailed methods for successful operation under a budget plan. From viewpoint of operating executives. 233 pp. $3.50.

Galloway. OFFICE MANAGEMENT. Analysis of conditions, standardization, control of personnel and production in departments of purchasing, sales, traffic, credit, advertising, accounting; labor-saving devices, wage plans. 701 pp. $3.00.

Galster. THE LABOR MOVEMENT IN THE SHOE INDUSTRY. Describes development and activities of national organization in shoe industry. 237 pp. $2.00.

Gardner. EFFECTIVE BUSINESS LETTERS. Definite instructions, ranging from details of form to general principles of structure and tone. 376 pp. $2.00.

Gardner. NEW COLLECTION METHODS. Retail,' department store, professional, wholesale, mail-order, and instalment collections. Credit and collection department organization, and procedure. Successful collection letters. 467 pp. $4.00.

Goldman. STOCK EXCHANGE LAW. Rights and duties of broker and customer. Significant cases cited. New York and Federal statutes, and Stock Exchange rules and regulations. 497 pp. $5.00.

Gowin. DEVELOPING EXECUTIVE ABILITY. How to plan the day's work effectively; practical short-cuts; mechanical aids; habits that develop initiative and reasoning power. 486 pp. $3.00.

Greeley. BUSINESS ACCOUNTING. A reading course prepared by leading accountants and educators for home study, offering working knowledge of principles, structure, and operation of accounts. Edited by Harold Dudley Greeley. Vol. I, "Theory of Accounts"; II, "Constructive Accounting"; III, "Cost Accounting"; IV, "Advanced and Analytical Accounting"; V, "Illustrative Accounting Problems"; "Reading Guide" outlining method of study. 2458 pp. $14.00.

Haney. THE BUSINESS OF RAILWAY TRANSPORTATION. A discussion of railway transportation as a business, mainly from the standpoint of the shipper. 613 pp. $4.00.

Himmelblau. COMPLETE ACCOUNTING COURSE. Prepared by members of faculty of Northwestern University. Edited by David Himmelblau, Professor of Accounting. Presentation combines questions, problems, and solutions with successive divisions of text material. Parts I, II, and III cover first, second, and third year's work respectively, each Part in 2 Units. Part I, Units I and II. $2.50 each. Practice sets for respective Units, $1.50 each. Publication dates and prices of Parts II and III to be announced.

Hudders. INDEXING AND FILING. A complete standard practice manual describing all details of filing work. 292 pp. $4.00.

Ignatius. THE FINANCING OF PUBLIC SERVICE CORPORATIONS. Part I, Public service corporations and regulation; Part II, Stock issues and dividends; Part III, Funded debt; Part IV, Capitalization. 508 pp. $5.00.

Ivey. PRINCIPLES OF MARKETING. Discusses basic problems, changing relations of manufacturers, wholesalers, and retailers, and present tendencies to combination and specialization. 351 pp. $3.00.

Jones (A. F.). LUMBER MANUFACTURING ACCOUNTS. From logging to sawed lumber. 112 pp. $3.00.

Jones (C. T.). TEACHING BUSINESS SUBJECTS IN THE SECONDARY SCHOOL. A survey by 18 leading business instructors. 307 pp. $2.50.

Jordan-Harris. COST ACCOUNTING PRINCIPLES AND PRACTICE. A standard work treating the entire flow of cost procedure and managerial uses of cost data. 550 pp. $4.00.

Keir. INDUSTRIAL ORGANIZATION. The natural resources of the United States, and their influence on its economic organization. A survey for students. 421 pp. $2.75.

Keir. MANUFACTURING INDUSTRIES IN AMERICA. The history of six basic industries—iron and steel, cotton, wool, leather, and paper—and a number of smaller ones. 324 pp. $3.00.

Kelly. TRAINING INDUSTRIAL WORKERS. Discusses national and state aid for vocational education; technical secondary schools; vestibule training; apprenticeship, training minor executives and foremen; transfers and promotion; factory surveys. 437 pp. $5.00.

Kennard. THE EDUCATIONAL DIRECTOR IN THE RETAIL STORE. Problems of employee training in modern stores. Work of the director as regards salespeople, the department, and the organization. 219 pp. $1.50.

Kennard. MERCHANDISE MANUALS For use in stores and schools, each volume covers a line of merchandise handled in retail stores—history, manufacture, classification, quality, and selling points. Individual volumes are: "Cotton and Linen," Thompson; "Glass & Glassware," Lehmann and Kennard; "Housefurnishing," Hutchinson; "Jewelry & Silverware," Kennard; "Leather Goods & Gloves," Lehmann; "Millinery," Aiken; "Notions," Souder; "Paper and Stationery," Lehmann and Kennard; "Silk," Thompson. 9 independent volumes, about 250 pp. each. $1.25 each.

Kester. ACCOUNTING: THEORY AND PRACTICE. A comprehensive treatment of the subject in 3 volumes. Vol. I—Covers fundamental principles, partnership and elementary corporation accounting. Vol. II—Covers valuation, study of balance sheet, depreciation, consolidations, liquidations. Vol. III—Shows how principles are applied to typical kinds of business. Vol. I, 625 pp., $3.00. Vol. II, 800 pp. $4.00. Vol. III, 718 pp., $5.00. Practice Blank Sets: accompanying Vol. I, $2.00; accompanying Vol. II, $1.75.

Konopak. COST ACCOUNTING FUNDAMENTALS. Introductory volume presenting cost accounting from management standpoint in perspective so that it can be grasped as a whole. 235 pp. $3.00.

Kooopman-Kester. FUNDAMENTALS OF ACCOUNTING. Thorough, high-school bookkeeping text, developed on accounting principles. Vol. I, first principles; Vol. II, controlling accounts; partnerships; corporations; and interrelations of balance sheet and profit and loss statement. Vol. I, 496 pp., $1.75. (Practice journal and ledger to accompany 36 cents.) Vol. II, 525 pp., $1.75. (Practice set to accompany, $1.95.)

Langston. PRACTICAL BANK OPERATION. Working manual describing in full detail all operations of banking as carried out in the National City Bank of New York. 2 vols., 713 pp. $8.00.

Langston-Whitney. BANKING PRACTICE. Textbook for colleges, based on "Practical Bank Operation" by Langston. 395 pp. $3.00.

Lee. BUSINESS WRITING. See Clapp Lee "Language for Men of Affairs."

Lewis. GETTING THE MOST OUT OF BUSINESS. Methods and policies of great business leaders and organizations. A philosophy of business inspiring and sound in practice. 515 pp. $3.00.

Lichtner. PLANNED CONTROL IN MANUFACTURING. Procedure for coordinating the work of a manufacturing enterprise through comprehensive planning. 329 pp. $5.00.

Lichtner. TIME STUDY AND JOB ANALYSIS. What executives want to know about standardizing working methods as a basis for control of manufacturing. 397 pp. $6.00.

Lough. BUSINESS FINANCE. A discussion of money and credit. Fundamentals of financial management; promotion; floating security issues; finance of going concerns; involvements and reorganizations. 631 pp. $3.00.

Lunt. SURETY BONDS. Clear analysis of nature of surety instruments, classes of risks, and considerations governing surety company. 370 pp. $2.50.

McAdow. MERCANTILE CREDITS. Emphasis on the credit man's skill and training rather than on department routine. 200 pp. $2.00.

McKinsey. BUDGETARY CONTROL. Gives full working procedure for building up final reports necessary for positive executive control. 474 pp. $4.25.

McKinsey-Meech. CONTROLLING THE FINANCES OF A BUSINESS. Includes determination of capital requirements; sources and devices for securing capital; disbursement of capital and distribution of profits. 638 pp. $5.00.

May. STREET RAILWAY ACCOUNTING. Applicable to either large system or small company, in conformity with Interstate Commerce Commission requirements. 454 pp. $5.00.

Maynard. MARKETING NORTHWESTERN APPLES. The methods which successfully market a highly specialized crop. 190 pp. $2.50.

Maynard-Weidler. AN INTRODUCTION TO BUSINESS. A survey course sketching the interrelation of business institutions. 600 pp. $4.00.

Meeker. THE WORK OF THE STOCK EXCHANGE. Investment transactions, margin trading, short selling, floor trader and specialist, odd-lot business, bond market, security collateral loan market, and Stock Exchange Clearing Corporation. 633 pp. $5.00.

Miles. ENGLISH IN BUSINESS. Textbook giving the essentials of high-school English, letter writing, oral reports, telephone conversation, sales talk, and advertising. 449 pp. $1.50.

Montgomery. AUDITING: THEORY AND PRACTICE. Vol. I—Foundation principles, full procedure in balance sheet or detailed audits, and preparation of statements, certificates and reports on completion. A standard work on business analysis. Vol. II—Specific procedure in audits of over 60 typical lines of business. Vol. I, 730 pp. $6.00. Vol. II, 559 pp., $4.00.

Montgomery. INCOME TAX PROCEDURE MANUALS. Annual publication covering the preparation returns item by item, considering law, accounting procedure, treasury rulings, and court decisions, and then giving specific advice. 9th annual edition, 1925. 1898 pp. $12.00. Editions back to 1918 available. Prices on request.

Montgomery-Staub. AUDITING PRINCIPLES. Textbook for colleges based on Montgomery's "Auditing Theory and Practice." 494 pp. $3.50.

Mucklow. REAL ESTATE ACCOUNTS. The special accounting methods required by brokers, property owners, or land development companies. 357 pp. $5.00.

Naylor. TRADE ASSOCIATIONS. History and fundamental theory; co-operation and competition; fair prices; law relating to associations; routine; commercial and industrial functions; duties of secretary. 389 pp. $5.00.

Nicholson-Rohrbach. COST ACCOUNTING. A practitioner's manual covering fundamental systems; factory routine; depreciation; maintenance; overhead; production. 576 pp. $5.00.

Nystrom. THE ECONOMICS OF RETAILING. Discusses consumption; expenses; salespeople; location and rent; prices; department stores; chain stores; mail-order houses; failure rates and public regulation. 404 pp. $3.00.

Olson - Hallman. CREDIT MANAGEMENT. A thorough treatment of the subject of credits and collections, with special emphasis on their relation to management. 444 pp. $4.50.

Palmer-Bell. ACCOUNTANTS' WORKING PAPERS. All schedules, analyses, and papers required in the ordinary audit, with many required under special conditions. 194 pp. $5.00.

Paton. ACCOUNTING THEORY. Restatement for modern corporate enterprise in terms of "properties" and "equities" without reference to "debit," "credit," or clerical routine. 508 pp. $4.00.

Radford. THE CONTROL OF QUALITY IN MANUFACTURING. Practical methods of maintaining positive quality control and thereby securing lower costs. 404 pp. $5.00.

Sakolski. THE PRINCIPLES OF INVESTMENT. A practical discussion by an analyst of experience covering the whole field, treating commercial paper, real estate mortgages, bonds and stock, as well as the machinery of investment. 565 pp. $4.50.

Saliers (Editor). ACCOUNTANTS' HANDBOOK. Reference manual giving in highly condensed form basic information covering entire range of accepted practice and related subjects. 33 sections—law, mathematics, corporations, financial statements, valuation, depreciation, finance, credits, etc. 1675 pp. $7.50.

Saliers. DEPRECIATION: PRINCIPLES AND APPLICATIONS. Exhaustive discussion from the accounting viewpoint, as regards both balance sheet valuation and operating requirements. 590 pp. $5.00.

Savage. INDUSTRIAL UNIONISM IN AMERICA. A study of the trend of labor organization. 344 pp. $2.25.

Savay. PRINCIPLES OF FOREIGN TRADE. Law; policies; government machinery; training representatives, technicalities of packing, shipping, customs, insurance and collection. 495 pp. $4.00

Scholfield. FILING DEPARTMENT OPERATION AND CONTROL. Comprehensive study of the problems relating to filing from the viewpoint of the responsible executive. 318 pp. $3.00.

Scott-Hayes. SCIENCE AND COMMON SENSE IN WORKING WITH MEN. Science as an aid in employment. Discusses tests of capacity and skill; workers' motives; and creation of opportunities. 154 pp. $2.00.

Scovell. INTEREST AS A COST. Critically examines every argument for and against the treatment of interest on investment as a cost. 254 pp. $2.50.

Shefferman. EMPLOYMENT METHODS. Deals with practical work of employment manager in factory, office, or store. 573 pp. $5.00.

Simonds. BUSINESS FUNDAMENTALS. Methods by which American leaders in finance and business have conducted their enterprises, formulated in 100 principles. 221 pp. $2.25.

Simons. PERSONNEL RELATIONS IN INDUSTRY. An introductory survey and guide-book to study in its field. 341 pp. $3.00.

Sprague. EXTENDED BOND TABLES. Bond yields and prices, normal and effective rates. Carried out to 8 decimal places. 233 pp. $10.00.

Sprague. THE PHILOSOPHY OF ACCOUNTS. An accounting classic dealing with the logic of the subject. 183 pp. $2.50.

Sprague-Perrine. THE ACCOUNTANCY OF INVESTMENT. All calculations for annuities, sinking funds, simple and compound interest, bond valuations, amortization schedules, loans, mortgages, etc., with illustrative problems. 371 pp. $5.00.

Stockwell. HOW TO READ A FINANCIAL STATEMENT. Deals with balance sheets as actually issued by business firms and corporations. Adapted especially to needs of credit men, bankers and investors. 443 pp. $4.50.

Taylor (F. M.). PRINCIPLES OF ECONOMICS. The fundamental theory of the subject. 587 pp. $3.00.

Taylor (P. S.). THE SAILORS' UNION OF THE PACIFIC. Based on records of government bodies, ship-owners' associations, and unions. 188 pp. $2.50.

Tipper. HUMAN FACTORS IN INDUSTRY. Discussion of background, psychology, and conditions of present-day industrial society. 280 pp. $2.00.

Tipper - Hotchkiss - Hollingworth - Parsons. THE PRINCIPLES OF ADVERTISING. A comprehensive work covering advertising economics; market analysis; planning campaigns; choosing media; preparing copy; designing display; organization of advertising work. 480 pp. $4.50.

Todman. WALL STREET ACCOUNTING. Accounting and auditing procedure of stock, cotton, and produce exchange brokers. 352 pp. $6.00.

Trottman. HISTORY OF THE UNION PACIFIC. The first transcontinental railroad—a vivid and detailed account of its construction, exploitation, receivership, and rehabilitation. 412 pp. $5.00.

Wade. SCALES AND WEIGHING: THEIR INDUSTRIAL APPLICATIONS. Weighing machinery of all types, especially as applied to business and manufacturing operations. 473 pp. $6.00.

Walter. THE RETAIL CHARGE ACCOUNT. Credit methods of leading retail stores described by the heads of 16 credit departments. 264 pp. $3.00.

Walton - Finney. MATHEMATICS OF ACCOUNTING AND FINANCE. Deals with more advanced problems of business than the ordinary business arithmetic. Covers all mathematics required by accountant and financial man. 274 pp. $4.00.

Ward. AMERICAN COMMERCIAL CREDITS. The commercial letter of credit, from legal and financial standpoints. 278 pp. $2.50.

Westerfield. BANKING PRINCIPLES AND PRACTICE. A standard treatise covering the entire subject. Modern banking theory and practice explained in detail; money and credit; national and state systems; internal bank organization and operation; foreign exchange. 5 volumes. 1370 pp. $12.00.

Westerfield. BANKING PRINCIPLES AND PRACTICE. A college text based upon the author's larger work (see above). 849 pp. $4.50.

Whitehead. PRINCIPLES OF SALESMANSHIP. A practical manual. The selling talk; the approach and interview; winning good-will; cultivating the characteristics of the successful salesman. 378 pp. $2.50.

Wilder - Buell. PUBLICITY. Comprehensive discussion of the means by which organizations or movements can win favorable public attention. 271 pp. $2.75.

Wilkes - York - Terrill. ELEMENTARY TRAINING FOR BUSINESS. Textbook for high schools and continuation schools, covering in sequence the work of typical office positions. Extensive practice material. 353 pp. $1.60. Practice set, $1.80.

Willis. THE FEDERAL RESERVE SYSTEM. Complete first-hand account. Passage of the Act; organization of the system; political and other influences; critical study of workings to date. 1765 pp. $10.00.

Willis-Byers. PORTLAND CEMENT PRICES. Detailed analysis of what determines current prices and price levels in the Portland cement industry. 123 pp. $1.25.

Woodbridge. ELEMENTS OF ACCOUNTING. A college text for students taking only one year of the subject, stressing its relations to economics. 450 pp. $3.50.

Woods. UNIFIED ACCOUNTING METHODS FOR INDUSTRIALS. Full working details for bringing accounting figures of manufacturing concerns into well-organized, summarized form enabling positive control. 484 pp. $6.00.

York. INTERNATIONAL EXCHANGE. Gives thorough grounding in principles and illustrates in detail practical methods of foreign trade financing. 600 pp. $5.00.

Young. STATISTICS AS APPLIED IN BUSINESS. A thorough discussion of statistical methods, giving the procedure for building up and directing a statistical department in a business concern. 675 pp. $5.00.

RONOGRAPH LIBRARY

Handy reference manuals, 5 x 7 inches in size averaging about 100 pages in length. Each volume treats intensively one specific topic or common problem of business practice, thus supplementing the text covering the field in question. $1.25 each, cash with order. The titles now available are:

Ackerman-Neuner—Credit Insurance (No. 12)
Altman—Cash Records (No. 20)
Altman—Recording Sales Transactions (No. 23)
Bartlett—Employment Records (No. 30)
Beach—Twenty Twenty-Minute Lessons in Bookkeeping (No. 19)
Beardsley—The Circular Advertising Department (No. 22)
Blackman—Business Mail (No. 3)
Brewster—Analyzing Credit Risks (No. 6)
Brewster—Bankruptcy (No. 29)
Brown—Statistical Typewriting (No. 14)
Dulin—Collection Letters (No. 2)
Dulin—Credit Letters (No. 1)
Fitting—Report Writing (No. 9)
Flaherty—How to Use the Dictionary (No. 18)
Gaines—The Art of Investment (No. 15)
Galloway—Organizing the Stenographic Department (No. 8)
Galston—Security Syndicate Operations (No. 33)
Giles—500 Answers to Sales Objections (No. 5)
Greeley—Estate Accounting (No. 24)
Griffiss—The New York Call Money Market (No. 32)
Hallman—Organizing the Credit Department (No. 27)
Kester—Depreciation (No. 28)
Lord—A Plan for Self-Management (No. 34)
McHale — Spanish and English Commercial Vocabulary (No. 36)
Nicholson—Profitable Management (No. 10)
Parker—Office Etiquette for Business Women (No. 31)
Patterson—Borrowing from Your Bank (No. 17)
Pinkerton—Accounting for Surplus (No. 21)
Sakolski—Elements of Bond Investment (No. 7)
Scammell—Use of the Telephone in Business (No. 25)
Scott—Influencing Men in Business (No. 13)
Shidle—Finding Your Job (No. 11)
Swindell—Newspaper Accounting (No. 16)
Waldron—Simplified Factory Accounting and Routing (No. 35)
Wallace—Filing Methods (No. 4)

THE RONALD PRESS COMPANY
15 East 26th Street, New York, N. Y.

311
Y84

UNIV OF MAINE
3 7550 01013609 2